Information encoded in volumes and areas of dendritic spines is nearly maximal across mammalian brains

Jan Karbowski[1,*], Paulina Urban[2,3,4]

[1] *Institute of Applied Mathematics and Mechanics, University of Warsaw, Warsaw, Poland;*

[2] *Laboratory of Functional and Structural Genomics, Centre of New Technologies, University of Warsaw, Warsaw, Poland;* [3] *College of Inter-Faculty Individual Studies in Mathematics and Natural Sciences, University of Warsaw, Warsaw, Poland;* [4] *Laboratory of Databases and Business Analytics, National Information Processing Institute, National Research Institute, Warsaw, Poland.*

Abstract

Long-term information associated with neuronal memory resides in dendritic spines. However, spines can have a limited size due to metabolic and neuroanatomical constraints, which should effectively limit the amount of encoded information in excitatory synapses. This study investigates how much information can be stored in the sizes of dendritic spines, and whether is it optimal in any sense? It is shown here, using empirical data for several mammalian brains across different regions and physiological conditions, that dendritic spines nearly maximize entropy contained in their volumes and surface areas for a given mean size in cortical and hippocampal regions. This result is essentially independent of the type of a fitting distribution to size data, as both short- and heavy-tailed distributions yield similar $90-100\%$ information efficiency in the majority of cases. On average, the highest information is contained in spine length ($4-5$ bits), and the lowest in spine volume and area ($2-3$ bits), although the latter two are closer to optimality. In contrast, we find that entropy density (entropy per spine size) is always suboptimal. Our results suggest a universality of entropy maximization in spine volumes and areas, which can be a new principle of memory storing in cortical and hippocampal synapses.

Keywords: Dendritic spine size distributions, Information encoded in spines; Synaptic memory efficiency; Entropy maximization; Constraints.

* Corresponding author: jkarbowski@mimuw.edu.pl

Significance statement

It is believed that memory in the brain is stored in the parts of excitatory synapses called dendritic spines. But how efficient is the memory capacity given synaptic size variability? Generally, the larger the synapse the more information can be packed in its structure. However, this process comes at some cost, as larger synapses use more metabolic energy and brain tissue. Thus, there exists a benefit-cost trade-off for storing long-term information in dendritic spines, with a unique optimal solution set by an average spine size. We show that volumes and areas of actual spines in different parts of the brain across different mammals and conditions are in majority of cases very close to that optimal theoretical solution for storing information, which is a remarkable result.

INTRODUCTION

Dendritic spines in excitatory neurons are the storing sites of long-term memory, mainly in their molecular components, i.e., receptors and proteins (Kennedy 2000; Kasai et al 2003; Chklovskii et al 2004; Lisman et al 2012; Huganir and Nicoll 2013; Meyer et al 2014; Takeuchi et al 2014). Experimental data show that spines vary vastly in sizes and shapes (Bonhoeffer and Yuste 2002; Holtmaat et al 2005; Yasumatsu et al 2008), but their size highly correlates with a magnitude of synaptic current (synaptic weight or strength), which suggests that there is a close relationship between spine structure and physiological function (Kasai et al 2003). Moreover, the variability in spine's size is also strictly associated with a variability in the number of AMPA receptors on spine membrane (Nusser et al 1998), as well as with changeability in the size of postsynaptic density (PSD) (Harris et al 1992), which is composed of the thousands of proteins implicated in molecular learning and memory storage (Lisman et al 2012; Sheng and Hoogenraad 2007).

The tight correlations between spine geometry and spine molecular composition, responsible for encoding and maintaining of memory, suggest that one can use the spine size as some measure of its information content. On the other hand, despite spine size variability, its mean value should be restricted due to limitations of cortical space associated with dense packing of different neuronal and non-neuronal components (Chklovskii et al 2002; Karbowski 2015). Moreover, because of the linear relationship between average spine size and postsynaptic current (Kasai et al 2003), bigger spines with stronger synaptic strength are generally more energy consuming than smaller spines with weaker strength (Attwell and Laughlin 2001; Karbowski 2012 and 2021; Laughlin et al 1998). These considerations suggest that molecular information encoded in the geometry of dendritic spines must be limited both by neuroanatomical and metabolic constraints, in particular, by the average spine size.

The main goal of this study is to investigate the information capacity of dendritic spines contained

in the distribution of their sizes. The specific questions we ask are the following: Is such information optimized in some sense, given the constraints on spine sizes? If so, how large is the deviation from the optimality for the parameters characterizing spine distributions? To answer these questions, we collected data from published literature on spine (or PSD) sizes (volumes, areas, length, and head width) for divers mammals and different cortical and subcortical regions (see the Methods). These data allowed us to compute empirical Shannon entropy associated with spine sizes for species, brain region and condition, and to compare it with a theoretical upper bound on the entropy for a given mean spine size.

RESULTS

Fitting of dendritic spine sizes to lognormal, loglogistic, and gamma distributions, and empirical Shannon entropy.

Previous empirical studies on dendritic spines have shown that their sizes (or synaptic weights) can be fitted well to either lognormal or gamma distributions (Benavides-Piccione et al 2013; Loewenstein et al 2011; Song et al 2005; Santuy et al 2020; Tamada et al 2020). However, these two types of the probability distributions differ significantly in terms of their asymptotic behavior: for very large spine sizes the former displays a heavy tail, while the latter decays exponentially with a short tail. Our first goal is to determine which distribution, with heavy or short tail, can better describe experimental data. In our analysis, we also added an additional probability density, not tried by previous studies, the loglogistic distribution that has a heavy tail. This distribution has an interesting property, because for small sizes it behaves similar to the gamma distribution, whereas for large sizes it resembles the lognormal distribution. For this reason, the loglogistic function is an intermediate alternative to the two extreme choices used in the past (Benavides-Piccione et al 2013; Loewenstein et al 2011; Song et al 2005; Santuy et al 2020; Tamada et al 2020).

In Figs. 1 and 2 we present histograms for the empirical spine volume and length from human cingulate cortex of two individuals (40 and 85 years old), taken from Benavides-Piccione et al (2013). These histograms can be well fitted to the three mentioned theoretical distributions. The goodness of these fits is conducted using the Kolmogorov-Smirnov test with cumulative distribution function CDF, which confirms that all three distributions can be used (Figs. 3 and 4). The quantitative agreement between the theoretical CDF and the empirical CDF depends on the number of bins N_b used for sampling, but in all studied cases the fits to the three theoretical distributions are always statistically significant at 95% level of confidence (Table 1). For example, for spine volume the best fits, as specified

by the Kolmogorov-Smirnov distance D_{KS}, are for lognormal and gamma distributions, but the latter distribution is a better choice for the larger N_b (Table 1). Similarly, for spine length the best fits are mostly for gamma distributions, except in one case where lognormal is slightly better (for maximal possible $N_b = 412$). Nevertheless, it should be said that the differences between D_{KS} for these two dominant distributions are rather small, which suggests that it may be difficult to precisely pin-point which distribution, either with short or heavy tail, is a better fit.

Next, we want to determine how stable are the distributions of spine sizes across the whole developmental period (Figs. 5-8). To address this, we use another collection of data on spine length and spine head diameter from human hippocampus (Das et al 2019; and private comm.). It is important to emphasize that spine length histograms seem visually invariant from the early age of 2 years to 71 years old, with the exception of 5 month old for which the histogram is broader though with essentially the same mean (Fig. 5). These empirical distributions of spine length can be significantly described (at 95% level of confidence) by all three theoretical distributions, however, the best fits (with the smallest D_{KS}) are provided by gamma distribution in almost all cases except for 5 months old, for which loglogistic is the best fit for maximal possible $N_b = 256$ (Fig. 6; Table 2). Essentially, this result agrees with the other fits of spine length from cingulate cortex present in Figs. 3 and 4, for which gamma distribution is also slightly superior.

The case for spine head diameter is slightly more complex, despite a visual similarity of all data histograms throughout the whole development (Fig. 7). Again, for this parameter, all three distributions yield statistically significant fits, but the best fit depends on the number of sampling bins N_b (Fig. 8; Table 2). Specifically, for small N_b the dominant distribution is the short-tailed gamma distribution, while for larger N_b the dominant is long-tailed loglogistic distribution.

Our next goal is to find empirical Shannon entropies associated with the discrete histograms of spine

sizes in Figs. 1, 2, 5, 7, and to compare them to continuous entropies associated with the three fitting probability distributions. This is conducted using Eq. 1 in the Methods for the continuous fitting distributions, and using a discrete version of this equation for the histograms. Both entropies depend on intrinsic noise amplitude Δx related to subspine molecular fluctuations (Eq. 2). Calculations reveal that the continuous distributions provide theoretical entropies (H_{th}) that very well approximate the empirical entropies (H_{em}) from the histograms for all fits in Figs. 1, 2, 5, and 7, with small differences between the two mostly no more than 5% (Table 3). Overall, the gamma distribution provides the best approximations, which in many cases deviate by less than 1% from the empirical entropies. The heavy-tailed lognormal distribution also gives very good approximations (mostly $1-3\%$), while the loglogistic distribution yields a little less accurate numbers but still sufficiently close (mostly $4-5\%$ of deviation).

To summarize, our fitting analysis reveals that the empirical data on spine sizes can be described well by the three different continuous distributions, either with short or heavy tails (gamma, lognormal, and loglogistic). Furthermore, the entropies associated with these distributions give reasonably good approximations to the empirical entropies.

Information contained in spine sizes is nearly maximal for spine volume and area. We collected a large data set of dendritic spine sizes from different sources (Tables 4-7; and Methods). These data contain values of mean and standard deviations of spine (or PSD) volume, surface area, length, and spine head diameter in different brain regions of several mammalian species (mostly cerebral cortex and hippocampus). Next, we make an assumption, based on the results and conclusions in Figs. 1-8 and Tables 1-3, that all the collected data on spine sizes can be described well by the three discussed above distributions, and additionally that the entropies associated with these continuous distributions are good approximations of the empirical entropies related to empirical, mostly unknown, distributions

of the collected data in Tables 4-7. With this logic in mind, we are able to estimate the entropy associated with each spine size parameter for the three distributions (Eq. 1 in the Methods), using only the means and standard deviations of the data. We denote the values of these continuous entropies H_g, H_{ln}, H_{ll} (Eqs. 6, 11, 17), and they provide a good measure of the empirical information content in dendritic spines. Additionally, we compute a theoretical upper bound on the entropy for a given mean spine size for each distribution ($H_{g,m}, H_{ln,m}, H_{ll,m}$; Eqs. 9, 14, 22). By comparing data driven continuous entropies H (H_g, H_{ln}, H_{ll}) to their corresponding upper bounds H_m ($H_{g,m}$, $H_{ln,m}$, $H_{ll,m}$), we can assess how close to the theoretical optimum is the information contained in spine sizes, which is quantified by an efficiency of information η defined as the ratio of H and H_m (Eq. 23 in the Methods). It is also good to keep in mind that the value $H_{g,m}$ is the maximal possible entropy for all existing probability distributions (not only for the three considered here) for a fixed mean spine size (e.g. Rieke et al 1999).

Another relevant quantity that we compute is deviation from the optimality D (D_g, D_{ln}, D_{ll}), which measures a deviation of the two parameters characterizing a given distribution from their optimal values defining a maximal entropy (Eqs. 31-33 in the Methods). This quantity is analogous to a standard error, which means its smallness is an indicator of the closeness of the empirical parameters to their optimal values.

The results in Figs. 9 and 10, and Tables 4 and 5, for spine volume and area (including PSD volumes and areas) indicate that the corresponding data driven continuous entropies essentially reach their upper theoretical bounds in a huge majority of cases. This fact is also evident from $90 - 100\%$ of the information efficiency ratio η, as well as from a relatively small values of the deviation from optimality D (Tables 4 and 5). There are, however, a few exceptions that yield suboptimal entropies, most notably cerebellum. There are also some negative values of entropy for gamma distributions

(macaque monkey), but these cases are only mathematical artifacts (see the Discussion).

The near optimality of entropy for spine volume and area across different species and many cortical and hippocampal regions is a remarkable result. It suggests that regardless of brain size, brain region, age, or neurophysiological condition, the distributions of spine volume and area adjust themselves such that to maximize their information content in most cases. The maximal values of the entropy depend logarithmically on the mean spine volume or area, and are in the range $2.3 - 3.5$ bits per spine, depending on average spine size.

The results for spine length and spine head diameter show that the corresponding entropies are more distant to their theoretical optima, especially for head diameter, than those for spine volume or area (lower values of the efficiency η and higher deviations D in Tables 6 and 7). Nevertheless, the spine length entropies, although not at the vicinity of the upper bound, are significantly closer to it than spine head diameter entropies (Figs. 11 and 12; Tables 6 and 7). Interestingly, the information contained in spine length and head diameter ($\sim 3 - 5$ bits) is greater than in spine volume and area.

It is important to emphasize that the above general trends for entropy maximization across various mammals are also consistent with the results for human brain, for which we do know the detailed probability distributions of spine sizes (Figs. 1-8). Specifically, for human cingulate cortex, the best fits to spine volumes are (mostly) gamma and lognormal distributions (Table 1), which yield entropy efficiency η respectively at the level of $97 - 99\%$ and $95 - 96\%$, with relatively low deviations from optimality D at $37 - 61\%$ and $20 - 24\%$ (Table 4). In contrast, the same cortical region for spine length with similar best fitting distributions gives a noticeable lower entropy efficiency about $86 - 90\%$, and much higher deviations in the range $184 - 478\%$ (Table 6). For human hippocampus, spine length is the best described by gamma distribution (Table 2), which yields comparable results: the entropy efficiency $88 - 90\%$ with deviations $273 - 348\%$ (Table 6). On the other hand, for human hippocampal

spine head diameter, which is best fitted by both gamma and loglogistic distributions, we obtain even lower entropy efficiency $84-86\%$, and slightly higher deviations for gamma distribution at $344-410\%$ (Table 7).

Taken together, the results shown in Figs. 9-13 indicate that although less information is encoded in spine volume and area than spine length or diameter, the former encoding is much more efficient. This means that the information associated with volume and surface area is maximal possible for given mean values of spine volume and area, across different species and different conditions.

Density of information in spines is not optimal.

As an alternative to the problem of information maximization in dendritic spines, we consider also the possibility that not entropy but the entropy density could be maximized. In particular, we study the ratio F of continuous data driven entropy to mean spine size (volume, area, length, and diameter), as a relevant quantity for maximization (Eq. 34 in the Methods). From a theoretical point of view, for each of the three considered distributions the entropy density F exhibits a single maximum. The height of that maximum corresponds to the upper bound on the entropy density, and our goal is to investigate how much the data driven F differs from that theoretical upper bound.

In Fig. 13 we present the data driven values of the entropy density and compare them to the maximal values of F. These results show that the densities of entropy F are generally far lower than their upper bounds. Specifically, F is only $\sim 30-40\%$ of its maximal values for spine length and spine head diameter (Fig. 13 C,D). The ratio F for spine volume is more scattered, with minimal values as low as 10% and maximal values up to $70-80\%$ of the maximal F (Fig. 13 A). The closest to the optimality is entropy density for spine area, but even here F is only in the range of $60-80\%$ of the upper bounds (Fig. 13 B).

We conclude that the density of entropy is not optimized in dendritic spines, which suggests that

this quantity is not the most relevant for information encoding in synapses.

Discussion

Summary of the main results: optimization of information encoded in spine volume and area, but not in spine length or head diameter.

Our results suggest that the sizes of dendritic spines can be described equally well by three different skewed distributions (gamma, lognormal, loglogistic), since all of them yield statistically significant fits. For the human data we used, this conclusion is independent on the number of bins, as verified by a goodness of fit procedure. Consequently, we calculated and compared spine entropies (i.e. information content in spines) for all these three probability densities, and confirmed that these continuous entropies are close to the the empirical entropies calculated from the histograms of the data sizes. This positive result allowed us to extend with certain confidence the entropy calculations also to the cases of other brain regions and species for which we do not know the empirical distributions of spine sizes, but rather only their means and standard deviations.

The main result of this study is that information encoded in volumes and areas of cortical and hippocampal dendritic spines is almost optimized, achieving in many cases a nearly 100% of its upper bound value for a given mean size of spine volume or area (Tables 4 and 5; Figs. 9 and 10). This conclusion is also supported by relatively small deviations from the optimality of the parameters characterizing the distributions (D values mostly below $20 - 30\%$ for heavy-tailed distributions). Furthermore, the optimality of information in spine volume or area is mostly independent of the investigated species (i.e. brain size), brain region (except cerebellum), age, or condition (Tables 4 and 5), which might suggest its possible universality, at least in the cerebral cortex and hippocampus. The exception of the cerebellum, with only $35 - 45\%$ of its information efficiency is intriguing, but from a theoretical point of view it is a consequence of the uniformity of spine sizes (though not of PSD sizes) in that part of the brain (Harris and Stevens 1988). That uniformity might serve some functional role,

different from information storing.

In most cases, the optimality of information in spine volumes and areas is also independent of the type of the chosen spine size distribution. Several exceptions for the gamma distribution in Tables 4 and 5 reflect the fact that this probability density poorly handles cases in which standard deviation exceeds the mean. That is also the reason for several negative entropy values associated with gamma distribution in Tables 4 and 5, which is only a mathematical artifact (see the Methods after Eq. 8), and relates to the fact that continuous distributions can yield negative entropies (in contrast to discrete distributions). Generally, the negative entropy cases appear almost exclusively for macaque monkey, which leads us to a hypothesis that spine volumes and areas in this species are probably better described by the heavy tailed distributions. Such distributions give information efficiency close to optimality (Tables 4 and 5).

The efficiency of the information encoded in spine length and spine head (or PSD) diameter is significantly lower than that encoded in spine volume and area (compare Tables 6 and 7 with Tables 4 and 5). Despite the relatively high efficiency values η for spine length (around $80-90\%$ in many cases), the deviations from optimality D are generally much larger than for spine volume or area (Fig. 11). For the spine head diameter the information efficiency η is generally the lowest and more scattered, mostly in the range $50-90\%$ (Fig. 12), and the deviations from optimality are even higher than for spine length (especially for gamma distribution). Again, in both cases the cerebellum is an outlier, although not so large as in the case of spine length. Taken together, these results suggest that neither spine length nor spine head diameter maintain the synaptic information optimally.

It is also shown here that the density of information is not optimized in any way, as its empirical values are far below their upper bounds (Fig. 13). This indicates that entropy density is not the relevant quantity for neuroanatomical and functional organization of dendritic spines, if we take an

evolutionary point of view into account (Parker and Maynard-Smith 1990; Striedter 2005).

Finally, there have been a few studies that are similar in spirit to our general approach but that differ in context and details. Notably, Brunel et al (2004) and Varshney et al (2004) assumed an information maximization under constraints to derive optimal distribution of synaptic weights and to make some general qualitative observations. In contrast, we took a different approach and we actually proved the near maximization of information in synaptic volumes and areas in different parts of the brain across several mammals. This was achieved using the collected data and fitting procedures for different distributions of synaptic sizes.

Stability of information contained in spine sizes and memory.

The aim of this study was to provide answers to two basic questions associated with the relationship between dendritic spines and memory. The first is: in which synaptic structural parameter information is most likely encoded? The second question is about the optimality of the encoded information. Both of the questions are interrelated.

We found that a typical dendritic spine can store between $1.9 - 3.4$ bits of information in its volume and a similar number in its surface area (Tables 4 and 5). For example, for spine volume in rat hippocampus we obtain $2.7 - 2.8$ bits, which is slightly smaller than the numbers calculated by Bartol et al (2015), $\sim 4 - 5$ bits, and the difference can be attributed to a different method of information estimation. The values of encoded bits in our study depend to some extent on the intrinsic noise amplitude Δx (Eq. 1), which is assumed to be set by relatively fast spatial fluctuations of actin molecules that provide structural backbone of dendritic spines and their plasticity (Matus 2000; Honkura et al 2008). Our calculations show that the noise amplitude Δx depends both on average spine size and on length ξ of an elementary unit of the actin polymer (called G-actin) (Eqs. 26,29, 30). In accordance with the empirical data we took $\xi = 0.007$ μm (Lodish et al 2016), but it is important

to underline that the entropy results are not too sensitive on the precise value of ξ. For example, information efficiency η for spine volume generally changes only by less than 1% if we modify the length ξ by as much as $\sim 30\%$. Our results indicate that spines in the mammalian cerebral cortex can hold about $(1-1.5)\cdot 10^{12}$ bits/cm^3 of memory in their volumes and areas, assuming in agreement with the data that cortical synaptic density is roughly brain size independent and on average $5\cdot 10^{11}$ cm^{-3} (Karbowski 2014; Sherwood et al 2020).

Although synaptic information can in principle be stored in many different structural synaptic parameters, it seems that the most relevant are those parameters that allow the maximal possible storage. This is what one can expect based on the principle of evolutionary pressure for optimizing the relevant variables (Parker and Maynard-Smith 1990; Striedter 2005). In this sense, and considering our results, the most likely parameter in which memory can be encoded is spine volume or area, and these parameters can serve as useful empirical indicators of a precise information content in individual synapses. Moreover, given that excitatory postsynaptic current and the number of AMPA receptors ("synaptic weight") are proportional to spine volume (Kasai et al 2003), this conclusion is consistent with a general hypothesis held in neuroscience that memory is encoded in excitatory synaptic weights.

Why can memory be more efficiently encoded in spine volume or area than e.g. spine length or diameter? It seems that volume offers more "states", as it intrinsically contains more degrees of freedom than any other linear size. Additionally, the fact that spine volume and area are equally good for storing information can be a consequence of their tight correlations (Bartol et al 2015). On the other hand, very weak correlations between spine length and spine volume (Bartol et al 2015; Tamada et al 2020) can explain the suboptimality of information storage in spine length (Fig. 11).

It is also curious that information efficiency η is relatively stable even in many cases when the condition changes. For example, for situations stress vs. non-stress, or mutation vs. control, or LTP

vs. non-LTP condition, the information efficiency stays remarkable constant despite changes in the mean and standard deviations of spine sizes (Tables 4-7). However, there are some exceptions to that stability. For example, the condition of stress can decrease the amount of stored information in volumes and areas of spines in the rat prefrontal cortex (data of Radley et al 2008; Tables 4 and 5), but paradoxically that same stress can increase the information encoded in spine length (Table 6). Similarly, in human prefrontal cortex, alzheimer's disease can decrease the stored information in spine length and head diameter (data of Boros et al 2019; Tables 6 and 7). Overall, the widespread stability of information efficiency might suggest that some compensatory mechanisms take plays in synapses that counteract a memory degradation. Interestingly, the same conclusion can be reached for the developmental data on human hippocampus (Das et al 2019), where the corresponding information encoded in spine length and head diameter is remarkably invariant across the human lifespan (Tables 6 and 7).

Implications for neuroanatomical and metabolic organization of the cerebral cortex in the context of synaptic information capacity.

It is known from physics that information is always associated with energy (Leff and Rex 1990). There have been suggestions that information processing in neurons is energy efficient (Laughlin et al 1998; Levy and Baxter 1996; Levy and Calvert 2021; Rieke et al 1999). We have an analogous situation in this study for dendritic spines: Long-term information contained in spine volumes and areas is maximized for a given mean spine size, which is directly proportional to the spine energy consumption during synaptic transmission (Attwell and Laughlin 2001; Karbowski 2012; Laughlin et al 1998). In this light, our result is essentially an additional example of the energetic efficiency of information, this time on a synaptic level and on a long time scale.

The nearly optimal encoding of information in dendritic spine volume and area, found in this

study, could be a new principle of neuroanatomical and functional organization of synaptic system in the mammalian cerebral cortex. Maximization of stored information in these synaptic parameters is largely independent of the cortical region and physiological conditions, and this suggests an universality of optimal encoding in synapses. In contrast, numerical spine density along dendrites generally lacks cross-regional uniformity, as it is greater in prefrontal than in somatosensory or visual cortical areas (Bianchi et al 2013; Elston et al 2006). Consequently, it seems that spine information content is decoupled from a linear spine dendritic density.

Another prominent principle, indirectly related, is associated with optimization (minimization) of neuronal wiring length or volume (Chklovskii et al 2002; Kaiser and Hilgetag 2006; Karbowski 2001; Mitchison 1991). It was shown that minimization of axonal and dendritic processes leads to an efficient use of metabolic energy in the brain (Attwell and Laughlin 2001; Chklovskii et al 2002), especially when the brain is scaled up in the evolutionary process (Herculano-Houzel 2011; Karbowski 2007; Yu et al 2014). A recent, interesting result shows that the entropy associated with the length of neuronal wiring is maximized (Song et al 2021). This result is complementary to ours, and both results suggest an efficient encoding of long-term information in the whole neuron. This conclusion is in line with the observation of optimal design of dendritic arbors for information content maximization (Wen et al 2009).

An alternative to neural wiring optimization, or its extension, is the principle of spine economy, which can explain the hierarchical organization of different components in the mammalian cerebral cortex (Karbowski 2015). It seems that it would be interesting to explore a connection of the spine economy principle with our results on the optimization of information in dendritic spines. It might lead to some new theoretical insight on the links between neuroanatomy and information content in cortical circuits, e.g. (Shepherd et al 2005; Bourne and Harris 2008; Hill et al 2012).

METHODS

Data collection and analysis for the sizes of dendritic spines.

All the experimental data on the sizes of dendritic spines used in this study were collected from different published sources, and concern several mammals. Data for mouse brain come from: (Arellano et al 2007; Benavides-Piccione et al 2002; Ishii et al 2018; Kashiwagi et al 2019; Konur et al 2003; Parajuli et al 2020; Rodriguez-Moreno et al 2018; Santuy et al 2020; Scheuss and Bonhoeffer 2014; Schikorski and Steven 1999; Tamada et al 2020), for rat brain from (Borczyk et al 2019; Cheetham et al 2014; Harris and Stevens 1988, 1989; Radley et al 2008; Swulius et al 2010), for rabbit brain from (Schwartzkroin and Kunkel 1982), for echidna brain from (Hassiotis et al 2003), for cat brain from (Clemo et al 2016; da Costa 2013), for macaque monkey brain from (Lubke et al 2015; Medalla and Barbas 2009; Medalla and Luebke 2015; Motley et al 2018; Young et al 2014), for dolphin brain from (Glezer and Morgane 1990), and for human brain from (Alonso-Nanclares et al 2008; Benavides-Piccione et al 2002; Benavides-Piccione et al 2013; Boros et al 2019; Das et al 2019; Tang et al 2001). The data used are both from cortical and subcortical regions, at different animal ages and different physiological conditions. The cortical areas include: piriform, somatosensory, visual, prefrontal, cingulate, parietal, temporal, and auditory cortex. The subcortical regions include: hippocampus, striatum, cerebellum, and forebrain.

The data for fitting distributions of dendritic spine sizes were taken from Benavides-Piccione et al (2013) (human cingulate cortex) and from Das et al (2019) (human hippocampus). For cingulate cortex we had either 5334 or 3355 data points related to spine volume, and either 5494 or 3469 data points related to spine length (for 40 and 85 years old individuals, respectively). For hippocampus we had the same number of data points for spine length and spine head diameter, but the numbers are

age-specific. In particular, we had 819 data points for 5 month old, 1108 for 2 yrs old, 915 for 23 yrs old, 657 for 27 yrs old, 1024 for 38 yrs old, 1052 for 45 yrs old, 662 for 57 yrs old, 805 for 58 yrs old, 1022 for 68 yrs old, 1026 for 70 yrs old, and 1032 for 71 yrs old. Histograms from these data were fitted to the three distributions (gamma, lognormal, and loglogistic) using standard tools from Matlab and Python. As a measure of the goodness of fit we used the Kolmogorov-Smirnov test (Keeping 1995; James 2006). For this, we constructed a cumulative distribution function CDF for each data set (empirical CDF) and compared it to three theoretical CDF corresponding to lognormal, loglogistic, and gamma probability densities with the same mean and variance as the original raw data. The test consists in determining a Kolmogorov-Smirnov distance, i.e., the maximum absolute deviation D_{KS} of the empirical CDF from the theoretical CDF, and then comparing such D_{KS} with a critical value of deviation $D_{N_b}(P)$, which depends on the number of sampling bins N_b and a level of significance P. If $D_{KS} < D_{N_b}(P)$ for a given N_b and significance P, then the fit is statistically significant (Keeping 1995). We used the level of significance $P = 0.05$ (corresponding to the confidence level of 95%), and a variable number of bins N_b, either 20, 30, or a maximal possible number for a given sample (412 for cingulate cortex, and 256 or 100 for hippocampus). Among the three theoretical distributions to which the data were compared, we chose that with the smallest Kolmogorov-Smirnov distance D_{KS} as the best fit. That best fit often depends on N_b. The code for calculations is provided in the Supplementary Information.

Theoretical modeling.

Maximization of entropy.

Shannon Entropy or information content H, contained in the probability distribution $\rho(x)$ of some

random variable x ($0 \leq x \leq \infty$) is defined as (Dayan and Abbott 2001)

$$H = -\frac{1}{\ln(2)} \int_0^\infty dx \rho(x) \ln[\rho(x)\Delta x], \tag{1}$$

where Δx is the intrinsic noise amplitude (fast undetectable fluctuations) associated with the variable x. We consider x to be either spine volume, spine surface area (or PSD area), spine head diameter, or spine length. The parameter Δx sets the scale for the resolution of x, and it depends on average spine size S (either volume, area, or length) in the following way (the estimate is given below):

$$\Delta x = \begin{cases} 0.084 S^{1/2}, & S - \text{length } (\mu m) \\ 0.223 S^{3/4}, & S - \text{area } (\mu m^2) \\ 0.226 S^{5/6}, & S - \text{volume } (\mu m^3). \end{cases} \tag{2}$$

In what follows, we solve the following optimization problem: we want to maximize the entropy given a fixed mean value of x. Keeping fixed the mean value of x reflects the constraint coming from neuroanatomical and/or metabolical restrictions on spine size. Mathematically, this procedure is equivalent to the standard Lagrange optimization problem with the Lagrangian L defined as

$$L = H - \lambda(\langle x \rangle - S), \tag{3}$$

where $\langle x \rangle$ is the mean (average) of x over the distribution $\rho(x)$, S is the average dendritic spine size, and λ is the Lagrange multiplier. The resulting maximal entropy will be a function of spine size S (see below).

Entropy of dendritic spines with gamma distribution.

The gamma distribution of a random variable x (spine size) is skewed but it decays fast for large values of x. It is defined as

$$\rho_g(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}, \tag{4}$$

where $\rho_g(x)$ is the probability density of x, the parameters α and β are some positive numbers, and $\Gamma(\alpha)$ is the Gamma function.

The entropy for the gamma distribution H_g is found from Eqs. (1) and (4), which generates

$$H_g(\alpha, \beta) = -\frac{1}{\ln(2)} \left[\ln\left(\frac{\beta^\alpha \Delta x}{\Gamma(\alpha)} \right) - \beta \langle x \rangle_g + (\alpha - 1) \int_0^\infty dx \rho_g(x) \ln(x) \right]. \tag{5}$$

The integral on the right hand side can be evaluated explicitly with the help of the formula (Gradshteyn and Ryzhik 2007)

$\int_0^\infty dx\, x^{\alpha-1} e^{-\beta x} \ln(x) = \Gamma(\alpha)[\psi(\alpha) - \ln(\beta)]/\beta^\alpha$,

where $\psi(\alpha)$ is the Digamma function (Gradshteyn and Ryzhik 2007). Additionally, the average of x for this distribution is $\langle x \rangle_g = \alpha/\beta$. Combining these results we obtain the entropy H_g as

$$H_g(\alpha, \beta) = \frac{1}{\ln(2)} \left[\alpha + (1 - \alpha)\psi(\alpha) + \ln\left(\frac{\Gamma(\alpha)}{\beta \Delta x} \right) \right]. \tag{6}$$

The standard deviation for the gamma distribution (defined as $\sigma_g = \sqrt{\langle x^2 \rangle_g - \langle x \rangle_g^2}$) is $\sigma_g = \sqrt{\alpha}/\beta$. We can invert the relations for the mean and standard deviations to find the parameters α and β for given experimental values of $\langle x \rangle_g$ and σ_g. The result is:

$$\alpha = \left(\frac{\langle x \rangle_g}{\sigma_g}\right)^2 \tag{7}$$

$$\beta = \frac{\langle x \rangle_g}{\sigma_g^2}. \tag{8}$$

Note that the entropy $H_g(\alpha, \beta)$ can become negative for $\alpha \ll 1$, since in this limit the Digamma function behaves as $\psi(\alpha) \approx -1/\alpha$. This situation corresponds to the cases in which $\sigma_g \gg \langle x \rangle_g$, i.e. to the data points for which standard deviation is much greater than the mean.

We find the maximal entropy of the gamma distribution given a fixed $\langle x \rangle_g$ by solving the Lagrange optimization problem defined in Eq. (3). The optimal parameters α and β are found by setting $\partial L/\partial\alpha = 0$, $\partial L/\partial\beta = 0$, and $\partial L/\partial\lambda = 0$. As a result, their optimal values are $\alpha_0 = 1$, $\beta_0 = 1/S$, and $\lambda_0 = -1/S$. For this values, the maximal entropy of the gamma distribution is:

$$H_{g,m} = \ln\left(\frac{eS}{\Delta x}\right) / \ln(2), \tag{9}$$

where e is the Euler number ($e \approx 2.71$). Interestingly, $H_{g,m}$ is also the maximal possible entropy given a fixed mean for *all* distributions of x.

Entropy of dendritic spines with lognormal distribution.

The lognormal distribution of a random variable x is both skewed and has a heavy tail, and is defined as

$$\rho_{ln}(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\frac{(\ln x - \mu)^2}{2\sigma^2}\right], \tag{10}$$

where $\rho_{ln}(x)$ is the probability density of x, and μ, σ are some parameters ($\sigma > 0$).

The entropy of the lognormal distribution H_{ln} can be determined by combining Eqs. (1) and (10). Using the substitution $y = \ln(x)$ the corresponding integrals can be evaluated similarly as in the case of gaussian distribution, which is a standard procedure. As a result, the entropy of lognormal distribution takes the form

$$H_{ln}(\mu, \sigma) = \frac{1}{\ln(2)} \left[\frac{1}{2} + \mu + \ln \left(\frac{\sqrt{2\pi}\sigma}{\Delta x} \right) \right]. \tag{11}$$

The average of x for this distribution is $\langle x \rangle_{ln} = \exp(\mu + \sigma^2/2)$, and the standard deviation ($\sigma_{ln} = \sqrt{\langle x^2 \rangle_{ln} - \langle x \rangle_{ln}^2}$) is $\sigma_{ln} = \langle x \rangle_{ln} \sqrt{e^{\sigma^2} - 1}$. By inverting these relations, we find

$$\mu = \ln \left(\frac{\langle x \rangle_{ln}^2}{\sqrt{\langle x \rangle_{ln}^2 + \sigma_{ln}^2}} \right) \tag{12}$$

$$\sigma = \sqrt{\ln\left(1 + (\sigma_{ln}/\langle x \rangle_{ln})^2\right)}. \tag{13}$$

Equations (12-13) allow us to find the characteristic parameters μ and σ defining the lognormal distribution from the experimental values of mean and standard deviation for the variable x.

As before, we find the maximal entropy of the lognormal distribution given a fixed $\langle x \rangle_{ln}$ by solving the Lagrange optimization problem defined in Eq. (3). The optimal parameters μ and σ are found by setting $\partial L/\partial \mu = 0$, $\partial L/\partial \sigma = 0$, and $\partial L/\partial \lambda = 0$. Their optimal values are: $\mu_0 = -0.5 + \ln(S)$, $\sigma_0 = 1$, and $\lambda_0 = -1/S$. For this values, the maximal entropy of the lognormal distribution is:

$$H_{ln,m} = \ln\left(\frac{\sqrt{2\pi}S}{\Delta x}\right) / \ln(2), \tag{14}$$

and it is clear that $H_{ln,m} < H_{g,m}$.

Entropy of dendritic spines with loglogistic distribution.

The loglogistic distribution of a random variable x is visually similar to the lognormal distribution with heavy tail, except that it decays as power law for very large x. The loglogistic probability density ρ_{ll} is defined as

$$\rho_{ll}(x) = \frac{b}{a}\frac{(x/a)^{b-1}}{[1+(x/a)^b]^2}, \tag{15}$$

where a and b are some positive parameters. Note that for $x \gg a$ the probability density ρ_{ll} behaves asymptotically as $\rho(x)_{ll} \sim 1/x^{b+1}$, which is a much slower decay than for the gamma distribution.

The entropy of the loglogistic distribution H_{ll} can be determined by combining Eqs. (1) and (15). Consequently, we have

$$H_{ll}(a,b) = \frac{1}{\ln(2)}\left[\ln\left(\frac{a}{b\Delta x}\right) - (b-1)\int_0^\infty dx\ \rho_{ll}(x)\ln\left(\frac{x}{a}\right) + 2\int_0^\infty dx\ \rho_{ll}(x)\ln\left(1+\left(\frac{x}{a}\right)^b\right)\right]. \tag{16}$$

The first integral on the right hand side (without the prefactor $(b-1)$) is performed by the substitution $y = (x/a)^b$. This transforms that integral into

$(1/b)\int_0^\infty dy \frac{\ln(y)}{(1+y)^2}$,

which is equal to 0 (Gradshteyn and Ryzhik 2007). The second integral on the right hand side (without the prefactor 2) can be transformed, using the same substitution, into

$\int_0^\infty dy \frac{\ln(1+y)}{(1+y)^2}$,

which has a value equal to 1 (Gradshteyn and Ryzhik 2007). As a result, the entropy of loglogistic distribution takes the form

$$H_{ll}(a,b) = \frac{1}{\ln(2)} \left[2 + \ln\left(\frac{a}{b\Delta x} \right) \right]. \tag{17}$$

The mean $\langle x \rangle_{ll}$ and standard deviation σ_{ll} for this distribution both exist, i.e. they are finite, provided the parameter $b > 2$. In this case, we have

$$\langle x \rangle_{ll} = \frac{\pi a/b}{\sin(\pi/b)} \tag{18}$$

$$\sigma_{ll} = a\sqrt{\frac{2\pi/b}{\sin(2\pi/b)} - \frac{(\pi/b)^2}{\sin^2(\pi/b)}}. \tag{19}$$

The inverse relations between the parameters a, b and $\langle x \rangle_{ll}, \sigma_{ll}$ are given by

$$\tan(\pi/b) = \left(1 + \frac{\sigma_{ll}^2}{\langle x \rangle_{ll}^2} \right) \left(\frac{\pi}{b} \right) \tag{20}$$

$$a = \langle x \rangle_{ll} \left(1 + \frac{\sigma_{ll}^2}{\langle x \rangle_{ll}^2} \right) \cos(\pi b). \tag{21}$$

The first equation has to be solved numerically, for given empirical values of $\langle x \rangle_{ll}, \sigma_{ll}$. The second

equation determines a once the parameter b is known.

The maximal value of the entropy for loglogistic distribution given a fixed $\langle x \rangle_{ll}$ is found by solving the Lagrange optimization problem defined in Eq. (3). The optimal parameters a and b are found by setting $\partial L/\partial a = 0$, $\partial L/\partial b = 0$, and $\partial L/\partial \lambda = 0$. Their optimal values are $a_0 = 2S/\pi$, $b_0 = 2$, and $\lambda_0 = -1/S$. For this values, the maximal entropy of the loglogistic distribution is:

$$H_{ll,m} = \ln\left(\frac{e^2 S}{\pi \Delta x}\right) / \ln(2), \tag{22}$$

and we also have that $H_{ll,m} < H_{g,m}$, as expected.

Definition of information efficiency.

Information efficiency η is defined as the ratio of the continuous entropy (H_{ln}, H_{ll}, H_g) to the theoretical upper bound of entropy for a given spine size $(H_{ln,m}, H_{ll,m}, H_{g,m})$. Thus, for a specific probability distribution we have

$$\eta_\alpha = \frac{H_\alpha}{H_{\alpha,m}}, \tag{23}$$

where α denotes one of the distributions (either ln, ll, or g).

Estimation of the intrinsic noise amplitude Δx.

Here we provide a justification for Eq. (2), which appeared above. Empirical data indicate that the size and shape of a dendritic spine is directly controlled by cytoskeleton, which consists mainly of the polymer filaments called F-actin (Matus 2000). Polymer F-actin is composed of many small monomers

called G-actin, each with a characteristic size 7 nm (Lodish et al 2016). Considering other proteins in a dendritic spine, this polymer has a relatively fast global turnover rate, about $40 - 50$ sec (Star et al 2002; Honkura et al 2008), which suggests that the length of F-actin (addition or removal of monomers) can fluctuate fast due to local intrinsic mechanisms. The fast fluctuations in F-actin length cause rapid fluctuations in the size and shape of a dendritic spine (Honkura et al 2008; Obashi et al 2019). Therefore, we assume that the amplitude of F-actin length fluctuations sets the scale for the intrinsic noise amplitude in an individual spine size, which we denote as Δx. We consider 3 separate cases for Δx related to spine length/diameter (Δx_{1D}), spine area (Δx_{2D}), and spine volume (Δx_{3D}).

a) 1D case.

Let L be the length of the longest branch of F-actin polymer in a spine (linear size of the longest actin rod). The value of L is also approximately the length of a spine (Kommaddi et al 2018). If this polymer consists of N monomers (G-actin) along its longest branch, each of length $\xi = 0.007$ μm (Lodish et al 2016), then $L = N\xi$. The length L fluctuates due to fluctuations in N, and we make the simplest assumption that the variability of N is governed by the Poisson stochastic process with the intensity parameter ν (the stationary Poisson probability of polymer of length N is $P_N = e^{-\nu}\nu^N/N!$). This assumption is consistent with empirical distributions of F-actin length in dendritic spines (Nanguneri et al 2019), as well as with the basic models of polymer growth (Hill 1980; Hu and Othmer 2011). Consequently, we have for the mean and standard deviation of N at the steady-state the following relations: $\langle N \rangle = \nu$, and $\sqrt{\langle N^2 \rangle - \langle N \rangle^2} = \sqrt{\nu}$, which leads to fluctuations in length L characterized by

$$\langle L \rangle = \xi\nu \tag{24}$$

and

$$\sqrt{\langle L^2\rangle - \langle L\rangle^2} = \xi\sqrt{\nu}. \tag{25}$$

We identify the intrinsic noise amplitude Δx_{1D} in this 1D case with the standard deviation of L, i.e., $\Delta x_{1D} = \sqrt{\langle L^2\rangle - \langle L\rangle^2}$. The last step is to combine Eqs. (24) and (25) such that to remove the unknown parameter ν, after which we obtain

$$\Delta x_{1D} = \sqrt{\xi\langle L\rangle} = 0.084\sqrt{\langle L\rangle}, \tag{26}$$

where $\langle L\rangle$ is in μm. Note that the noise amplitude Δx_{1D} in 1D case is proportional to the square root of mean spine length $\langle L\rangle$.

b) 2D case.

We assume, in agreement with the data, that the spine surface area is dominated by the surface area of the spine head, which is approximately a sphere. Thus the spine area A is approximately $A = \pi D^2$, where D is both the spine head diameter and (in analogy to 1D case) the linear size of the largest branch of F-actin polymer in the spine head, i.e., we have $D = N\xi$. Furthermore, we have for the mean area $\langle A\rangle = \pi\xi^2\langle N^2\rangle$, and for the standard deviation of area $\sqrt{\langle A^2\rangle - \langle A\rangle^2} = \pi\xi^2\sqrt{\langle N^4\rangle - \langle N^2\rangle^2}$.

The moments of N for the Poisson distribution of N with the intensity parameter ν are: $\langle N^2\rangle = \nu^2 + \nu$, and $\langle N^4\rangle = \nu^4 + 6\nu^3 + 7\nu^2 + \nu$, and thus we have an equation for the unknown parameter ν:

$$\nu^2 + \nu = \frac{\langle A\rangle}{\pi\xi^2}. \tag{27}$$

Since the value of $\pi\xi^2$ is $1.5 \cdot 10^{-4}$ μm^2, and this is much smaller than any recorded value of spine (or PSD) area $\langle A \rangle$ in Table 5 (by a factor of at least 200), we can safely assume that the parameter $\nu \gg 1$. This implies that we can neglect the linear term on the left in Eq. (27), and obtain

$$\nu \approx \sqrt{\frac{\langle A \rangle}{\pi\xi^2}}. \tag{28}$$

Similarly, for the standard deviation of A we have $\sqrt{\langle A^2 \rangle - \langle A \rangle^2} = \pi\xi^2\sqrt{4\nu^3 + 6\nu^2 + \nu} \approx 2\pi\xi^2\nu^{3/2}$. We identify the intrinsic noise amplitude Δx_{2D} in this 2D case with the standard deviation of A, and thus obtain

$$\Delta x_{2D} \approx 2\pi^{1/4}\sqrt{\xi}\langle A \rangle^{3/4} = 0.223\langle A \rangle^{3/4}, \tag{29}$$

where the average spine area $\langle A \rangle$ is in μm^2. Note that the noise amplitude Δx_{2D} in 2D case is proportional to $\langle A \rangle^{3/4}$.

c) 3D case.

We make a similar assumption as in 2D case, that volume V of a spine is dominated by the volume of spine head, which is approximately a sphere. Repeating similar steps as before, we have for the mean volume

$\langle V \rangle = (\pi/6)\xi^3\langle N^3 \rangle = (\pi/6)\xi^3(\nu^3 + 3\nu^2 + \nu) \approx (\pi/6)\xi^3\nu^3,$

and for the standard deviation of volume

$\sqrt{\langle V^2 \rangle - \langle V \rangle^2} = (\pi/6)\xi^3\sqrt{\langle N^6 \rangle - \langle N^3 \rangle^2} = (\pi/6)\xi^3\sqrt{9\nu^5 + 54\nu^4 + 84\nu^3 + 30\nu^2 + \nu} \approx (\pi/2)\xi^3\nu^{5/2},$

where we used the fact that $\nu \gg 1$, and the following moments of the Poisson distribution: $\langle N^3 \rangle = \nu^3 + 3\nu^2 + \nu$, and $\langle N^6 \rangle = \nu^6 + 15\nu^5 + 65\nu^4 + 90\nu^3 + 31\nu^2 + \nu$.

Combining of the equations for the mean and standard deviation, and identifying the intrinsic noise amplitude Δx_{3D} in this 3D case with the volume standard deviation, we obtain

$$\Delta x_{3D} \approx \left(\frac{3^5\pi}{2}\right)^{1/6} \sqrt{\xi}\langle V \rangle^{5/6} = 0.226\langle V \rangle^{5/6}, \tag{30}$$

where the average spine volume $\langle V \rangle$ is in μm^3. Note that the noise amplitude Δx_{3D} in 3D case is nearly a linear function of mean spine volume. This is a very similar result to the empirical finding in Yasumatsu et al (2008), where it was found that the standard deviation of intrinsic spine fluctuations could be well fitted to a linear function of spine volume.

Deviation from optimality of the empirical parameters characterizing spine size distributions.

We introduce a measure of deviation of the parameters characterizing a given distribution from their optimal theoretical values, as a relative combined error from optimality.

For the gamma distribution the deviation D_g is:

$$D_g = \frac{1}{\sqrt{2}}\sqrt{\left(\frac{\alpha - \alpha_0}{\alpha_0}\right)^2 + \left(\frac{\beta - \beta_0}{\beta_0}\right)^2}. \tag{31}$$

For the log-normal distribution the deviation D_{ln} is:

$$D_{ln} = \frac{1}{\sqrt{2}}\sqrt{\left(\frac{\mu - \mu_0}{\mu_0}\right)^2 + \left(\frac{\sigma - \sigma_0}{\sigma_0}\right)^2}, \tag{32}$$

and for the log-logistic distribution the deviation D_{ll} is:

$$D_{ll} = \frac{1}{\sqrt{2}}\sqrt{\left(\frac{a-a_0}{a_0}\right)^2 + \left(\frac{b-b_0}{b_0}\right)^2}. \tag{33}$$

In Eqs. (31-33) the parameters $\alpha_0, \beta_0, \mu_0, \sigma_0$, and a_0, b_0 are the optimal parameters for a given distribution. The smaller the value of D, the closer the empirical distribution is to its maximal entropy. For example, if both α and β deviate from their respective optimal values by 50%, then $D_g = 1/2$, or the deviation is 50% in Tables 4-7. Similarly for the rest of the parameters.

Alternative measure of optimality: Maximization of entropy density.

As an alternative to the above approach with the maximization of the entropy for a given average spine size, we consider below the maximization of the density of entropy. More precisely, we analyze the optimization problem in which we want to maximize the entropy of spine sizes per average spine size.

The fitness function F in this case takes the form:

$$F = \frac{H}{\langle x \rangle}, \tag{34}$$

where $\langle x \rangle$ is the mean spine size as above. Interestingly, the function F displays maxima for each of the three probability densities.

For the gamma distribution, we have the following form of the fitness function $F_g(\alpha, \beta) = H_g/\langle x \rangle_g$:

$$F_g(\alpha, \beta) = \frac{\beta}{\alpha \ln(2)}\left[\alpha + (1-\alpha)\psi(\alpha) + \ln\left(\frac{\Gamma(\alpha)}{\beta \Delta x}\right)\right]. \tag{35}$$

Optimal values of α and β are found from the conditions: $\partial F_g(\alpha,\beta)/\partial\alpha = 0$, and $\partial F_g(\alpha,\beta)/\partial\beta = 0$. A straightforward calculation leads to the following optimal values: $\alpha_0 = 1$, and $\beta_0 = 1/\Delta x$, for which the maximal value of fitness function $F_{g,m} \equiv F_g(\alpha_0,\beta_0)$ is

$$F_{g,m} = \frac{1}{\Delta x \ln(2)}. \tag{36}$$

For the lognormal distribution, we have the following form of the fitness function $F_{ln}(\mu,\sigma) = H_{ln}/\langle x\rangle_{ln}$:

$$F_{ln}(\mu,\sigma) = \frac{1}{\ln(2)}\left[\frac{1}{2} + \ln\left(\frac{\sqrt{2\pi}\sigma}{\Delta x}\right) + \mu\right] e^{-(\mu+\sigma^2/2)}. \tag{37}$$

The optimal values μ_0, σ_0 of μ and σ are found from the conditions: $\partial F_{ln}(\mu,\sigma)/\partial\mu = 0$, and $\partial F_{ln}(\mu,\sigma)/\partial\sigma = 0$. This procedure yields: $\mu_0 = 1/2 + \ln(\Delta x/\sqrt{2\pi})$ and $\sigma_0 = 1$, and the corresponding maximal value of entropy density $F_{ln,m} = F_{ln}(\mu_0,\sigma_0)$, which is

$$F_{ln,m} = \frac{\sqrt{2\pi}}{e\Delta x \ln(2)}. \tag{38}$$

For the loglogistic distribution, we have the following form of the fitness function $F_{ll}(a,b) = H_{ll}/\langle x\rangle_{ll}$:

$$F_{ll}(a,b) = \frac{b}{\pi a \ln(2)} \sin\left(\frac{\pi}{b}\right) \ln\left(\frac{e^2 a}{b\Delta x}\right). \tag{39}$$

As before maximization of $F_{ll}(a,b)$ is performed by solving the following conditions: $\partial F_{ll}(a,b)/\partial a = 0$, and $\partial F_{ll}(a,b)/\partial b = 0$. This procedure yields the optimal parameters of a and b: $a_0 = 2\Delta x/e$ and $b_0 = 2$, and the maximal value of F_{ll} as

$$F_{ll,m} = \frac{e}{\pi \Delta x \ln(2)}. \tag{40}$$

Note that generally we have the following inequalities: $F_{g,m} > F_{ln,m} > F_{ll,m}$, which means that gamma distribution provides the highest maximal information density.

Acknowledgments

We would like to thank Dr. Ruth Benavides-Piccione and Dr. Sujan Chandra Das for providing us the datasets of their published results on dendritic spine sizes in humans. The work was supported by the Polish National Science Centre (NCN) grant number 2021/41/B/ST3/04300 (JK).

Competing interests

There is no competing interests to declare.

Ethic statement

Ethic statements does not apply to this study, because all experimental data were collected from other sources.

References

Alonso-Nanclares L, Gonzalez-Sorlano J, Rodriguez JR, DeFelipe J (2008) Gender differences in human cortical synaptic density. *Proc. Natl. Acad. Sci. USA* **105**: 14615-14619.

Arellano JI, Benavides-Piccione R, DeFelipe J, Yuste R (2007) Ultrastructure of dendritic spines: correlation between synaptic and spine morphologies. *Front. Neurosci.* **1**: 131-143.

Attwell D, Laughlin SB (2001) An energy budget for signaling in the gray matter of the brain. *J. Cereb. Blood Flow Metabol.* **21**: 1133-1145.

Bartol TM, Bromer C, Kinney J, Chirillo MA, Bourne JN, Harris KM, Sejnowski TJ (2015) Nanoconnectomic upper bound on the variability of synaptic plasticity. *eLife* **4**: e10778.

Benavides-Piccione R, Ballesteros-Yanez I, DeFelipe J, Yuste R (2002) Cortical area and species differences in dendritic spine morphology. *J. Neurocytol.* **31**: 337-346.

Benavides-Piccione R, Fernaud-Espinosa I, Robles V, Yuste R, DeFelipe J (2013) Age-based comparison of human dendritic spine structure using complete three-dimensional reconstructions. *Cerebral Cortex* **23**: 1798-1810.

Bianchi S, Stimpson CD, Bauernfeind AL, Schapiro SJ, Baze WB, et al (2013) Dendritic morphology of pyramidal neurons in the chimpanzee neocortex: Regional specializations and comparison to humans. *Cereb. Cortex* **23**: 2429-2436.

Bonhoeffer T, Yuste R (2002) Spine motility: phenomenology, mechanisms, and function. *Neuron* **35**: 1019-1027.

Borczyk M, Sliwinska MA, Caly A, Bernas T, Radwanska K (2019) Neuronal plasticity affects correlation between the size of dendritic spine and its postsynaptic density. *Scientific Rep.* **9**: 1693.

Boros BD, Greathouse KM, Gearing M, Herskowitz JH (2019) Dendritic spine remodeling accompanies Alzheimer's disease pathology and genetic susceptibility in cognitively normal aging. *Neurobiology of*

Aging **73**: 92-103.

Bourne JN, Harris KM (2008) Balancing structure and function at hippocampal dendritic spines. *Annual Rev. Neurosci.* **31**: 47-67.

Brunel N, Hakim V, Isope P, Nadal JP, Barbour B (2004) Optimal information storage and the distribution of synaptic weights: perceptron versus Purkinje cell. *Neuron* **43**: 745-757.

Cheetham CEJ, Barnes SJ, Albieri G, Knott GW, Finnert GT (2014) Pansynaptic enlargement at adult cortical connections strengthened by experience. *Cereb. Cortex* **24**: 521-531.

Chklovskii DB, Mel BW, Svoboda K (2004) Cortical rewiring and information storage. *Nature* **431**: 782-788.

Chklovskii DB, Schikorski T, Stevens CF (2002) Wiring optimization in cortical circuits. *Neuron* **34**: 341-347.

Clemo HR, Lomber SG, Meredith MA (2016) Synaptic basis for cross-modal plasticity: enhanced supragranual dendritic spine density in anterior ectosylvian auditory cortex of the early deaf cat. *Cereb. Cortex* **26**: 1365-1376.

da Costa NM (2013) Diversity of thalamorecipient spine morphology in cat visual cortex and its implication for synaptic plasticity. *J. Comp. Neurol.* **521**: 2058-2066.

Das SC, Chen D, Callor WB, Christensen E, Coon H, Williams ME (2019) Dil-mediated analysis of presynaptic and postsynaptic structures in human postmortem brain tissue. *J. Comp. Neurol.* **527**: 3087-3098.

Dayan P, Abbott LF (2001) *Theoretical Neuroscience.* Cambridge, MA: MIT Press.

Elston GN, Benavides-Piccione R, Elston A, Zietsch B, DeFelipe J, et al (2006) Specializations of the granular prefrontal cortex of primates: Implications for cognitive processing. *Anatomical Record A* **288A**: 26-35.

Glezer II, Morgane PJ (1990) Ultrastructure of synapses and golgi analysis of neurons in neocortex of the lateral gyrus (visual cortex) of the dolphin and pilot whale. *Brain Res. Bulletin* **24**: 401-427.

Gradshteyn IS, Ryzhik IM (2007) *Table of Integrals, Series, and Products.* Elsevier: Amsterdam (17th edit.).

Harris KM, Stevens JK (1988) Dendrritic spines of rat cereberal purkinje cells: serial electron microscopy with reference to their biophysical characteristics. *J. Neurosci.* **8**: 4455-4469.

Harris KM, Stevens JK (1989) Dendrritic spines of CA1 pyramidal cells in the rat hippocampus: serial electron microscopy with reference to their biophysical characteristics. *J. Neurosci.* **9**: 2982-2997.

Harris KM et al (1992) Three-dimensional structure of dendritic spines and synapses in rat hippocampus (CA1) at postnatal day 15 and adult ages: implications for the maturation of synaptic physiology and long-term potentiation. *J. Neurosci.* **12**: 2685-2705.

Hassiotis M, Paxinos G, Ashwell KWS (2003) The anatomy of the cerebral cortex of the echidna (Tachyglossus aculeatus). *Comp. Biochem. Physiol.* **A 136**: 827-850.

Herculano-Houzel S (2011) Scaling of brain metabolism with a fixed energy budget per neuron: Implications for neuronal activity, plasticity and evolution. *PLoS ONE* **6**: e17514.

Hill SL, Wang Y, Riachi I, Schurmann F, Markram H (2012) Statistical connectivity provides a sufficient foundation for specific functional connectivity in neocortical neural microcircuits. *Proc. Natl. Acad. Sci. USA* **109**: E2885-E2894.

Hill TL (1980) Bioenergetic aspects and polymer length distribution in steady-state head-to-tail polymerization of actin or microtubules. *Proc. Natl. Acad. Sci. USA* **77**: 4803-4807.

Holtmaat AJ, Trachtenberg JT, Wilbrecht L, Shepherd GM, Zhang X, et al (2005) Transient and persistent dendritic spines in the neocortex in vivo. *Neuron* **45**: 279-291.

Honkura N, Matsuzaki M, Noguchi J, Ellis-Davies GCR, Kasai H (2008) The subspine organization of

actin fibers regulates the structure and plasticity of dendritic spines. *Neuron* **57**: 719-729.

Hu J, Othmer HG (2011) A theoretical analysis of filament length fluctuations in actin and other polymers. *J. Math. Biol.* **63**: 1001-1049.

Huganir RL, Nicoll RA (2013) AMPARs and synaptic plasticity: the last 25 years. *Neuron* **80**: 704-717.

Ishii K, Nagaoka A, Kishida Y, Okazaki H, Yagishita S, Ucar H, Takahashi N, Saito N, Kasai H (2018). In vivo dynamics of dendritic spines in the neocortex of wild-type and Fmr1 KO mice. *eNeuro* **5**: ENEURO.0282-18.2018.

James F (2006) *Statistical Methods in Experimental Physics.* London: World Scientific, 2 edition.

Kaiser M, Hilgetag CC (2006) Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput. Biol.* **2**: e95.

Karbowski J (2001) Optimal wiring principle and plateaus in the degree of separation for cortical neurons. *Phys. Rev. Lett.* **86**: 3674-3677.

Karbowski J (2007) Global and regional brain metabolic scaling and its functional consequences. *BMC Biology* **5**: 18.

Karbowski J (2012) Approximate invariance of metabolic energy per synapse during development in mammalian brains. *PLoS ONE* **7**: e33425.

Karbowski J (2014) Constancy and trade-offs in the neuroanatomical and metabolic design of the cerebral cortex. *Front. Neural Circuits* **8**: 9.

Karbowski J (2015) Cortical composition hierarchy driven by spine proportion economical maximization or wire volume minimization. *PloS Comput. Biol.* **11**: e1004532.

Karbowski J (2021) Energetics of stochastic BCM type synaptic plasticity and storing of accurate information. *J. Comput. Neurosci.* **49**: 71-106.

Kasai H, Matsuzaki M, Noguchi J, Yasumatsu N, Nakahara H (2003) Structure-stability-function relationships of dendritic spines. *Trends Neurosci.* **26**: 360-368.

Kashiwagi Y, Higashi T, Obashi K, Sato Y, Komiyama NH, Grant SGN, Okabe S (2019) Computational geometry analysis of dendritic spines by structured illumination microscopy. *Nature Comm.* **10**: 1285.

Keeping ES (1995) *Introduction to Statistical Inference.* New York: Dover. Kennedy MB (2000) Signal-processing machines at the postsynaptic density. *Science* **290**: 750-754.

Kommaddi RP et al (2018) Aβ mediates F-actin disassembly in dendritic spines leading to cognitive deficits in Alzheimer's disease. *J. Neurosci.* **38**: 1085-1099.

Konur S, Rabinowitz D, Fenstermaker V, Yuste R (2003) Systematic regulation of spine head diameters and densities in pyramidal neurons from juvenile mice. *J. Neurobiol.* **56**: 95-112.

Laughlin SB, de Ruyter van Steveninck RR, Anderson JC (1998) The metabolic cost of neural information. *Nature Neurosci.* **1**: 36-40.

Leff HS, Rex AF (1990) *Maxwell's Demon: Entropy, Information, Computing.* Princeton, NJ: Princeton Univ. Press.

Levy WB, Baxter RA (1996) Energy efficient neural codes. *Neural Computation* **8**: 531-543.

Levy WB, Calvert VG (2021) Communication consumes 35 times more energy than computation in the human cortex, but both costs are needed to predict synapse number. *Proc. Natl. Acad. Sci. USA* **118**: e2008173118.

Lisman J, Yasuda R, Raghavachari S (2012) Mechanisms of CaMKII action in long-term potentiation. *Nat. Rev. Neurosci.* **13**: 169-182.

Lodish HF, Berk A, Kaiser C, Krieger M, Bretscher A, Ploegh H, et al (2016) Cell organization and movement I: Microfilaments. In *Molecular Cell Biology.* New York: Freeman (8th edit.).

Loewenstein Y, Kuras A, Rumpel S (2011) Multiplicative dynamics underlie the emergence of the log-normal distribution of spine sizes in the neocortex in vivo. *J. Neurosci.* **31**: 9481-9488.

Luebke JI, Medalla M, Amatrudo JM, Weaver CM, Crimins JL, Hunt B, Hof PR, Peters A (2015) Age-related changes to layer 3 pyramidal cells in the rhesus monkey visual cortex. *Cereb. Cortex* **25**: 1454-1468.

Matus A (2000) Actin-based plasticity in dendritic spines. *Science* **290**: 754-758.

Medalla M, Barbas H (2009) Synapses with inhibitory neurons differentiate anterior cingulate from dorsolateral prefrontal pathways associated with cognitive control. *Neuron* **61**: 609-620.

Medalla M, Luebke JI (2015) Diversity of glutamatergic synaptic strength in lateral prefrontal versus primary visual cortices in the rhesus monkey. *J. Neurosci.* **35**: 112-127.

Meyer D, Bonhoeffer T, Scheuss V (2014) Balance and stability of synaptic structures during synaptic plasticity. *Neuron* **82**: 430-443.

Mitchison G (1991) Neuronal branching patterns and the economy of cortical wiring. *Proc. R. Soc. Lond. B* **245**: 151-158.

Motley SE, Grossman YS, Janssen WGM, Baxter MG, Rapp PR, Dumitriu D, Morrison JH (2018) Selective loss of thin spines in area 7a of the primate intraparietal sulcus predicts age-related working memory impairment. *J. Neurosci.* **38**: 10467-10478.

Nanguneri S, et al (2019) Characterization of nanoscale organization of F-actin in morphologically distinct dendritic spines in vitro using supervised learning. *eNeuro* 0425-18.2019, 1-13.

Nusser Z et al (1998) Cell type and pathway dependence of synaptic AMPA receptor number and variability in the hippocampus. *Neuron* **21**: 545-559.

Obashi K, Matsuda A, Inoue Y, Okabe S (2019) Precise temporal regulation of molecular diffusion within dendritic spines by actin polymers during structural plasticity. *Cell Reports* **27**: 1503-1515.

Parajuli LK, Wako K, Maruo S, Kakuta S, Taguchi T, Ikuno M, Yamakado H, Takahashi R, Koike M (2020) Developmental changes in dendritic spine morphology in the striatum and their alteration in an A53T α-synuclein transgenic mouse model of Parkinson's disease. *eNeuro* **7**: ENEURO.0072-20.2020.

Parker GA, Maynard Smith J (1990) Optimality theory in evolutionary biology. *Nature* **348**: 27-33.

Radley JJ, Rocher AB, Rodriguez A, Ehlenberger DB, Dammann M, Mcewen BS, Morrison JH, Wearne SL, Hof PR (2008) Repeated stress alters dendritic spine morphology in the rat medial prefrontal cortex. *J. Comp. Neurol.* **507**: 1141-1150.

Rieke F, Warland D, de Ruyter R, Bialek W (1999) *Spikes: Exploring the neural code.* Cambridge, MA: MIT Press.

Rodriguez-Moreno J, et al (2018) Quantitative 3D ultrastructure of thalamocortical synapses from the "lemniscal" ventral posteromedial nucleus in mouse barrel cortex. *Cereb. Cortex* **28**: 3159-3175.

Santuy A, et al (2020) Estimation of the number of synapses in the hippocampus and brain-wide by volume electron microscopy and genetic labeling. *Scientific Reports* **10**: 14014.

Scheuss V, Bonhoeffer T (2014) Function of dendritic spines on hippocampal inhibitory neurons. *Cereb. Cortex* **24**: 3142-3153.

Schikorski T, Stevens CF (1999) Quantitative fine-structural analysis of olfactory cortical synapses. *Proc. Natl. Acad. Sci. USA* **96**: 4107-4112.

Schwartzkroin PA, Kunkel DD (1982) Electrophysiology and morphology of the developing hippocampus of fetal rabbits. *J. Neurosci.* **2**: 448-462.

Sheng M, Hoogenraad CC (2007) The postsynaptic architecture of excitatory synapses: a more quantitative view. *Annu. Rev. Biochem.* **76**: 823-847.

Shepherd GM, Stepanyants A, Bureau I, Chklovskii D, Svoboda K (2005) Geometric and functional organization of cortical circuits. *Nat. Neurosci.* **8**: 782-790.

Sherwood CC, Miller SB, Karl M, Stimpson CD, Phillips KA, et al (2020) Invariant synapse density and neuronal connectivity scaling in primate neocortical evolution. *Cereb. Cortex* **30**: 5604-5615.

Song S, Sjöström PJ, Reigl M, Nelson SB, Chklovskii DB (2005) Highly nonrandom features of synaptic connectivity in local cortical circuits. *PLoS Biol.* **3**: e68.

Song Y, Zhou D, Li S (2021) Maximum entropy principle underlies wiring length distribution in brain networks. *Cereb. Cortex* **31**: 4628-4641.

Star EN, Kwiatkowski DJ, Murthy VN (2002) Rapid turnover of actin in dendritic spines and its regulation by activity. *Nature Neurosci.* **5**: 239-246.

Striedter GF (2005) *Principles of Brain Evolution.* Sunderland, MA: Sinauer Assoc.

Swulius MT, Kubota Y, Forest A, Waxham MN (2010) Structure and composition of the postsynaptic density during development. *J. Comp. Neurol.* **518**: 4243-4260.

Takeuchi T, Duszkiewicz AJ, Morris RGM (2014) The synaptic plasticity and memory hypothesis: encoding, storage and persistence. *Phil. Trans. R. Soc. B* **369**: 20130288.

Tamada H, Blanc J, Korogod N, Petersen CCH, Knott GW (2020) Ultrastructural comparison of dendritic spine morphology preserved with cryo and chemical fixation. *eLife* **9**: e56384.

Tang Y, Nyengaard JR, de Groot DMG, Gundersen HJG (2001) Total regional and global number of synapses in the human brain neocortex. *Synapse* **41**: 258-273.

Wen Q, Stepanyants A, Elston G, Grosberg AY, Chklovskii DB (2009) Maximization of the connectivity repertoire as a statistical principle governing the shapes of dendritic arbors. *Proc. Natl. Acad. Sci. USA* **106**: 12536-12541.

Varshney LR, Sjostrom PJ, Chklovskii DB (2006) Optimal information storage in noisy synapses under resource constraint. *Neuron* **52**: 409-423.

Yasumatsu N, Matsuzaki M, Miyazaki T, Noguchi J, Kasai H (2008) Principles of long-term dynamics

of dendritic spines. *J. Neurosci.* **28**: 13592-13608.

Young ME, Ohm DT, Dumitriu D, Rapp PR, Morrison JH (2014) Differential effects of aging on dendritic spines in visual cortex and prefrontal cortex of the rhesus monkey. *Neuroscience* **274**: 33-43.

Yu YG, Karbowski J, Sachdev RNS, Feng JF (2014) Effect of temperature and glia in brain size enlargement and origin of allometric body-brain size scaling in vertebrates. *BMC Evol. Biol.* **14**: 178.

Figure Captions

Fig. 1

Fitting of spine volume to lognormal, loglogistic, and gamma distributions. Empirical histograms of spine volumes from human cingulate cortex (rectangles; combined spines from apical and basal dendrites taken from Benavides-Piccione (2013)) were fitted to three different distributions (solid lines). Number of bins $N_b = 20$ for all plots. A) Fits for 40 years old yield the following parameters: $\mu = -1.56$, $\sigma = 0.87$ (lognormal); $a = 0.22$, $b = 2.03$ (loglogistic); and $\alpha = 1.67$, $\beta = 5.72$ (gamma). B) Fits for 85 years old give: $\mu = -1.50$, $\sigma = 0.91$ (lognormal); $a = 0.24$, $b = 1.93$ (loglogistic); and $\alpha = 1.56$, $\beta = 4.91$ (gamma).

Fig. 2

Fitting of spine length to lognormal, loglogistic, and gamma distributions. Empirical histograms of spine length from human cingulate cortex (rectangles; combined spines from apical and basal dendrites taken from Benavides-Piccione (2013)) were fitted to three different distributions (solid lines). Number of bins $N_b = 20$ for all plots. A) Fits for 40 years old yield: $\mu = 0.34$, $\sigma = 0.52$ (lognormal); $a = 1.46$, $b = 3.44$ (loglogistic); and $\alpha = 4.31$, $\beta = 2.71$ (gamma). B) Fits for 85 years old yield: $\mu = 0.33$, $\sigma = 0.52$ (lognormal); $a = 1.44$, $b = 3.42$ (loglogistic); and $\alpha = 4.26$, $\beta = 2.70$ (gamma).

Fig. 3

Cumulative distribution function CDF of spine volume for empirical and theoretical distributions in cingulate cortex. Note that all three theoretical CDF are close to the empirical CDF, but gamma and lognormal are slightly better than loglogistic.

Fig. 4

Similar as in Fig. 3, but for spine length.

Fig. 5

Similarity of spine length distributions across human lifespan. Empirical data for human hippocampal spine length (rectangles; taken from Das et al (2019)) ranging from infancy (except 5 month old), through maturity, to senility look very similar. These data were fitted to three different distributions (solid lines). A) Fitting parameters for lognormal: $\mu = 0.14$, $\sigma = 0.59$, (5 months); $\mu = 0.17$, $\sigma = 0.55$, (2 years); $\mu = 0.19$, $\sigma = 0.52$, (23 years); $\mu = 0.20$, $\sigma = 0.53$, (27 years); $\mu = 0.20$, $\sigma = 0.52$, (38 years); $\mu = 0.18$, $\sigma = 0.53$, (45 years); $\mu = 0.17$, $\sigma = 0.52$, (57 years); $\mu = 0.16$, $\sigma = 0.53$, (58 years); $\mu = 0.16$, $\sigma = 0.53$, (68 years); $\mu = 0.17$, $\sigma = 0.53$, (70 years); $\mu = 0.17$, $\sigma = 0.53$, (71 years).

B) Fitting parameters for loglogistic: $a = 1.20$, $b = 2.86$, (5 months); $a = 1.23$, $b = 3.13$, (2 years); $a = 1.26$, $b = 3.33$, (23 years); $a = 1.26$, $b = 3.23$, (27 years); $a = 1.26$, $b = 3.33$, (38 years); $a = 1.23$, $b = 3.23$, (45 years); $a = 1.22$, $b = 3.33$, (57 years); $a = 1.21$, $b = 3.23$, (58 years); $a = 1.21$, $b = 3.23$, (68 years); $a = 1.22$, $b = 3.23$, (70 years); $a = 1.21$, $b = 3.23$, (71 years).

C) Fitting parameters for gamma: $\alpha = 3.35$, $\beta = 2.48$, (5 months); $\alpha = 3.84$, $\beta = 2.83$, (2 years); $\alpha = 4.17$, $\beta = 3.03$, (23 years); $\alpha = 4.07$, $\beta = 2.93$, (27 years); $\alpha = 4.13$, $\beta = 2.99$, (38 years); $\alpha = 4.06$, $\beta = 2.99$, (45 years); $\alpha = 4.11$, $\beta = 3.05$, (57 years); $\alpha = 4.03$, $\beta = 3.03$, (58 years); $\alpha = 4.01$, $\beta = 2.99$, (68 years); $\alpha = 4.07$, $\beta = 3.02$, (70 years); $\alpha = 4.07$, $\beta = 3.04$, (71 years).

Fig. 6

Cumulative distribution function CDF of spine length for empirical and theoretical distributions in hippocampus. Note a near overlap of all three theoretical CDF with the empirical

CDF.

Fig. 7

Similarity of spine head diameter distributions across human lifespan. Empirical data for human hippocampal spine head diameter (rectangles; taken from Das et al (2019)) ranging from infancy, through maturity, to senility look very similar. These data were fitted to three different distributions (solid lines). A) Fitting parameters for lognormal: $\mu = -1.08$, $\sigma = 0.47$, (5 months); $\mu = -1.07$, $\sigma = 0.48$, (2 years); $\mu = -1.05$, $\sigma = 0.47$, (23 years); $\mu = -1.05$, $\sigma = 0.49$, (27 years); $\mu = -1.06$, $\sigma = 0.49$, (38 years); $\mu = -1.05$, $\sigma = 0.48$, (45 years); $\mu = -1.05$, $\sigma = 0.48$, (57 years); $\mu = -1.06$, $\sigma = 0.48$, (58 years); $\mu = -1.06$, $\sigma = 0.48$, (68 years); $\mu = -1.06$, $\sigma = 0.49$, (70 years); $\mu = -1.06$, $\sigma = 0.49$, (71 years).

B) Fitting parameters for loglogistic: $a = 0.35$, $b = 3.85$, (5 months); $a = 0.35$, $b = 3.70$, (2 years); $a = 0.35$, $b = 3.85$, (23 years); $a = 0.36$, $b = 3.70$, (27 years); $a = 0.35$, $b = 3.70$, (38 years); $a = 0.36$, $b = 3.70$, (45 years); $a = 0.36$, $b = 3.85$, (57 years); $a = 0.35$, $b = 3.85$, (58 years); $a = 0.35$, $b = 3.70$, (68 years); $a = 0.35$, $b = 3.70$, (70 years); $a = 0.35$, $b = 3.70$, (71 years).

C) Fitting parameters for gamma: $\alpha = 5.03$, $\beta = 13.44$, (5 months); $\alpha = 4.82$, $\beta = 12.68$, (2 years); $\alpha = 5.01$, $\beta = 12.98$, (23 years); $\alpha = 4.77$, $\beta = 12.19$, (27 years); $\alpha = 4.59$, $\beta = 11.80$, (38 years); $\alpha = 4.80$, $\beta = 12.27$, (45 years); $\alpha = 4.90$, $\beta = 12.55$, (57 years); $\alpha = 4.88$, $\beta = 12.68$, (58 years); $\alpha = 4.73$, $\beta = 12.21$, (68 years); $\alpha = 4.63$, $\beta = 11.96$, (70 years); $\alpha = 4.59$, $\beta = 11.79$, (71 years).

Fig. 8

Similar as in Fig. 6, but for spine head diameter.

Fig. 9

Information encoded in spine volume is nearly optimal across brains with different sizes, regions, and physiological conditions. Note a close alignment of the data points to the maximal entropy curve (solid line), especially for lognormal and loglogistic distributions. The gamma distribution has several outliers, which are rather mathematical artifacts. The visible outlier for all three distributions is the point corresponding to rat cerebellum. Legend for data points: diamonds for mouse, circles for rat, squares for macaque monkey, triangles for human.

Fig. 10

Information encoded in spine surface area is nearly optimal across mammalian brains with different regions and conditions. Similar as in Fig. 9, the nearly optimal alignment of data points to the maximal entropy curve (solid line), with a cerebellum as an outlier. Legend for data points: diamonds for mouse, circles for rat, pentagram for cat, squares for macaque monkey, triangles for human.

Fig. 11

Information encoded in spine length is suboptimal. Note that the data driven continuous entropies are generally suboptimal (especially for the gamma distribution), but nevertheless they are relatively close to the maximal entropy curve (solid line). Legend for data points: diamonds for mouse, circles for rat, pentagram for cat, squares for macaque monkey, triangles for human.

Fig. 12

Information encoded in spine head diameter is suboptimal. Note scattered data points, far away from the maximal entropy curve (solid line). Legend for data points: diamonds for mouse, circles for rat, stars for rabbit, plus for echidna, pentagram for cat, squares for macaque monkey, hexagon for

dolphin, triangles for human.

Fig. 13

Density of information stored in dendritic spines is far from optimal. Density of entropy F as a function of (A) spine volume, (B) spine surface area, (C) spine length, and (D) spine head diameter. Blue diamonds correspond to data points described by lognormal distribution, black squares correspond to loglogistic, and red triangles to gamma distribution. Note that all these data points are far below the theoretical upper bounds for the entropy density represented by three lines (blue solid line for lognormal, black dashed line for loglogistic, and green dotted line for gamma distribution).

Table 1: Kolmogorov-Smirnov goodness of fit for spine volume and length: human cingulate cortex.

Bin no N_b	Spine Parameter	Age (yrs)	Kolmogorov-Smirnov distance D_{KS}		
			Lognormal	Loglogistic	Gamma
20					
	Volume	40	0.269*	0.272	0.274
	Volume	85	0.256*	0.259	0.256*
	Length	40	0.190	0.199	0.179*
	Length	85	0.133	0.140	0.122*
30					
	Volume	40	0.221	0.230	0.212*
	Volume	85	0.193	0.201	0.183*
	Length	40	0.127	0.137	0.117*
	Length	85	0.091	0.094	0.080*
412					
	Volume	40	0.026	0.034	0.020*
	Volume	85	0.033	0.037	0.019*
	Length	40	0.019*	0.026	0.020
	Length	85	0.018	0.029	0.014*

All fits to the three theoretical distributions are statistically significant at the level of $P = 0.05$ (regardless of the number of bins N_b), since the Kolmogorov-Smirnov distance D_{KS} is always smaller than the critical distance D_{cr}, which is 0.294 for $N_b = 20$, 0.240 for $N_b = 30$, and 0.067 for $N_b = 412$ (Keeping 1995). The asterisk indicates the smallest KS distance, and hence the best theoretical distribution from the chosen three.

Table 2: Kolmogorov-Smirnov goodness of fit for spine length and head diameter: human hippocampus.

Bin N_b	Spine Param.	Distrib.	Kolmogorov-Smirnov distance D_{KS} Age (yrs)										
			$\frac{5}{12}$	2	23	27	38	45	57	58	68	70	71
20	Length	Lognormal	0.128	0.126	0.123	0.127	0.127	0.127	0.129	0.131	0.128	0.127	0.128
		Loglogistic	0.135	0.134	0.134	0.138	0.138	0.138	0.139	0.141	0.138	0.138	0.139
		Gamma	0.123*	0.118*	0.117*	0.122*	0.122*	0.122*	0.124*	0.127*	0.123*	0.122*	0.124*
20	H. diam.	Lognormal	0.120	0.131	0.135	0.159	0.162	0.163	0.165	0.169	0.166	0.198	0.196
		Loglogistic	0.119	0.130	0.135	0.157*	0.169	0.171	0.173	0.176	0.174	0.193*	0.191
		Gamma	0.116*	0.128*	0.133*	0.157*	0.159*	0.159*	0.161*	0.166*	0.163*	0.201	0.198*
30	Length	Lognormal	0.099	0.099	0.091	0.092	0.092	0.090	0.090	0.092	0.093	0.092	0.093
		Loglogistic	0.107	0.108	0.102	0.100	0.099	0.099	0.100	0.102	0.103	0.103	0.103
		Gamma	0.094*	0.092*	0.085*	0.084*	0.085*	0.085*	0.085*	0.087*	0.088*	0.088*	0.088*
30	H. diam.	Lognormal	0.087	0.098	0.098	0.113	0.123	0.126	0.129	0.131	0.128	0.135*	0.133*
		Loglogistic	0.080*	0.091*	0.088*	0.112	0.113*	0.116	0.119*	0.120*	0.117*	0.138	0.137
		Gamma	0.081	0.098	0.100	0.110*	0.127	0.130	0.132	0.135	0.132	0.140	0.138
256	Length	Lognormal	0.088	0.070	0.063	0.061	0.058	0.057	0.055	0.056	0.055	0.055	0.054
		Loglogistic	0.069*	0.049	0.038	0.038	0.038	0.035	0.035	0.036	0.036	0.037	0.036
		Gamma	0.077	0.046*	0.035*	0.036*	0.034*	0.035*	0.033*	0.036*	0.034*	0.033*	0.034*
100	H. diam.	Lognormal	0.062	0.057	0.055	0.055	0.046	0.047	0.047	0.048	0.047	0.045	0.044
		Loglogistic	0.032*	0.038*	0.038*	0.035*	0.031*	0.034*	0.035*	0.034*	0.034*	0.031*	0.031
		Gamma	0.033	0.054	0.055	0.050	0.048	0.050	0.051	0.053	0.051	0.049	0.046

All fits to the three theoretical distributions are statistically significant at the level of $P = 0.05$ (regardless of the number of bins N_b), since the Kolmogorov-Smirnov distance D_{KS} is always smaller than the critical distance D_{cr}, which is 0.294 for $N_b = 20$, 0.240 for $N_b = 30$, and 0.067 for $N_b = 412$ (Keeping 1995). The asterisk indicates the smallest KS distance, and hence the best theoretical distribution from the chosen three.

Table 3: Comparison of empirical entropies H_{em} with theoretical entropies H_{th}.

Spine Parameter	Age (yrs)	Empirical entropy H_{em} (bits)	Theoretical entropy H_{th} (bits)		
			Lognormal	Loglogistic	Gamma
		Cingulate Cortex:			
Volume	40	3.21	3.05 (4.9%)	2.88 (10.3%)	3.23 (0.6%)
	85	3.20	3.05 (4.7%)	2.89 (9.7%)	3.23 (0.9%)
Length	40	4.66	4.61 (1.1%)	4.51 (3.2%)	4.71 (1.1%)
	85	4.55	4.64 (2.0%)	4.53 (0.4%)	4.74 (4.2%)
		Hippocampus:			
Length	$\frac{5}{12}$	4.58	4.63 (1.1%)	4.52 (1.3%)	4.75 (3.7%)
	2	4.68	4.58 (2.1%)	4.47 (4.5%)	4.69 (0.2%)
	23	4.68	4.56 (2.6%)	4.45 (4.9%)	4.66 (0.4%)
	27	4.70	4.58 (2.6%)	4.48 (4.7%)	4.69 (0.2%)
	38	4.69	4.57 (2.6%)	4.47 (4.7%)	4.68 (0.2%)
	45	4.70	4.57 (2.8%)	4.47 (4.9%)	4.68 (0.4%)
	57	4.70	4.56 (3.0%)	4.46 (5.1%)	4.67 (0.6%)
	58	4.69	4.56 (2.8%)	4.46 (4.9%)	4.68 (0.2%)
	68	4.69	4.57 (2.6%)	4.47 (4.7%)	4.68 (0.2%)
	70	4.69	4.57 (2.6%)	4.46 (4.9%)	4.68 (0.2%)
	71	4.69	4.56 (2.8%)	4.46 (4.9%)	4.67 (0.4%)
Head diam.	$\frac{5}{12}$	3.63	3.55 (2.2%)	3.45 (5.0%)	3.64 (0.3%)
	2	3.71	3.60 (3.0%)	3.50 (5.7%)	3.70 (0.3%)
	23	3.67	3.59 (2.2%)	3.50 (4.6%)	3.69 (0.5%)
	27	3.67	3.63 (1.1%)	3.53 (3.8%)	3.73 (1.6%)
	38	3.74	3.65 (2.4%)	3.54 (5.3%)	3.75 (0.3%)
	45	3.72	3.63 (2.4%)	3.53 (5.1%)	3.73 (0.3%)
	57	3.71	3.61 (2.7%)	3.52 (5.1%)	3.72 (0.3%)
	58	3.71	3.61 (2.7%)	3.51 (5.4%)	3.71 (0%)
	68	3.73	3.63 (2.7%)	3.53 (5.4%)	3.73 (0%)
	70	3.68	3.64 (1.1%)	3.54 (3.8%)	3.75 (1.9%)
	71	3.71	3.65 (1.6%)	3.55 (4.3%)	3.76 (1.3%)

All the data for empirical entropies correspond to $N_b = 20$. For cingulate cortex, the data come from combining apical and basal dendrites. The numbers in the brackets for theoretical entropies correspond to differences between H_{em} and H_{th}, i.e., $|H_{th} - H_{em}|/H_{em}$, given in %.

Table 4: Information contained in the distribution of spine volumes.

Species/ region (cond.)	Spine volume mean±SD (μm^3)	Entropy (bits)			Efficiency % (Deviation %)		
		$H_{ln}(H_{ln,m})$	$H_{ll}(H_{ll,m})$	$H_g(H_{g,m})$	$\eta_{ln}(D_{ln})$	$\eta_{ll}(D_{ll})$	$\eta_g(D_g)$
Mouse							
neoctx (wt)	0.152±0.111[a]	2.81 (3.01)	2.66 (2.92)	2.99 (3.13)	93.4 (25.9)	91.0 (47.8)	95.5 (87.5)
neoctx (mut.)	0.144±0.131[a]	2.92 (3.00)	2.75 (2.91)	3.10 (3.11)	97.4 (16.8)	94.6 (34.5)	99.5 (20.8)
neoctx	0.081±0.087[b]	2.83 (2.86)	2.66 (2.76)	2.96 (2.97)	99.2 (9.2)	96.3 (26.6)	99.7 (13.3)
hippo.	0.076±0.082[c]	2.82 (2.84)	2.65 (2.75)	2.95 (2.96)	99.2 (8.9)	96.4 (26.4)	99.6 (14.1)
hippo.	0.038±0.036[d]	2.61 (2.67)	2.44 (2.58)	2.78 (2.79)	97.7 (14.5)	94.5 (32.4)	99.8 (11.4)
piriform 1a	0.129±0.174[d]	2.97 (2.97)	2.83 (2.88)	2.86 (3.09)	99.9 (1.4)	98.2 (18.3)	92.8 (45.0)
piriform 1b	0.071±0.073[d]	2.79 (2.82)	2.62 (2.73)	2.94 (2.94)	98.8 (11.1)	95.8 (28.5)	99.9 (5.4)
somatosen.	0.06±0.04[e]	2.52 (2.78)	2.37 (2.69)	2.68 (2.90)	90.5 (28.6)	88.2 (54.5)	92.4 (125.0)
visual	0.09±0.07[f]	2.72 (2.88)	2.56 (2.79)	2.90 (3.00)	94.5 (23.0)	91.8 (43.6)	96.8 (65.3)
striat. (wt, 1m)	0.050±0.075[g]	2.73 (2.74)	2.61 (2.65)	2.40 (2.85)	99.6 (6.3)	98.6 (15.3)	84.1 (55.6)
striat. (mut, 1m)	0.077±0.100[g]	2.84 (2.84)	2.69 (2.75)	2.80 (2.96)	100.0 (0.4)	97.9 (19.5)	94.5 (40.7)
striat. (wt, 3m)	0.074±0.094[g]	2.83 (2.83)	2.68 (2.74)	2.82 (2.95)	100.0 (1.5)	97.7 (20.3)	95.5 (38.0)
striat. (mut, 3m)	0.089±0.162[g]	2.82 (2.88)	2.77 (2.79)	1.78 (3.00)	97.9 (15.8)	99.3 (10.9)	59.5 (69.8)
striat. (wt, 6m)	0.084±0.115[g]	2.86 (2.86)	2.72 (2.77)	2.73 (2.98)	100.0 (2.1)	98.2 (17.8)	91.7 (46.6)
striat. (mut, 6m)	0.088±0.143[g]	2.85 (2.88)	2.76 (2.78)	2.29 (2.99)	99.1 (10.3)	99.0 (13.3)	76.4 (62.1)
striat. (wt, 22m)	0.143±0.211[g]	2.99 (2.99)	2.86 (2.90)	2.70 (3.11)	99.7 (5.8)	98.6 (15.7)	86.8 (54.1)
striat. (mut, 22m)	0.089±0.146[g]	2.85 (2.88)	2.76 (2.79)	2.25 (3.00)	99.0 (10.8)	99.0 (13.1)	75.2 (62.8)
Rat							
hippo.	0.062±0.080[h]	2.79 (2.79)	2.64 (2.70)	2.75 (2.91)	100.0 (0.7)	97.8 (19.7)	94.7 (39.9)
hippo. (ctr)	0.067±0.085[i]	2.81 (2.81)	2.66 (2.72)	2.79 (2.93)	99.9 (1.5)	97.7 (20.3)	95.5 (37.9)
hippo. (cLTP)	0.080±0.101[i]	2.85 (2.85)	2.70 (2.76)	2.84 (2.97)	99.9 (1.8)	97.7 (20.5)	95.8 (37.3)
prefro. (ctr)	0.09±0.18[j]	2.78 (2.88)	2.78 (2.79)	1.20 (3.00)	96.7 (20.4)	99.5 (9.1)	40.0 (75.0)
prefro. (stress)	0.08±0.37 [j]	2.15 (2.85)	2.76 (2.76)	-19.71 (2.97)	75.4 (59.3)	99.9 (1.8)	- (95.3)
somatosen. (ctr)	0.51±0.35[k]	3.06 (3.31)	2.92 (3.22)	3.23 (3.42)	92.6 (32.6)	90.7 (52.3)	94.4 (112.3)
somatosen. (LTP)	0.79±0.49[k]	3.09 (3.41)	2.96 (3.32)	3.24 (3.53)	90.5 (44.4)	89.0 (60.4)	91.8 (159.9)
cerebel.	0.12±0.02[l]	1.06 (2.95)	1.02 (2.86)	1.08 (3.07)	35.9 (60.4)	35.8 (322.7)	35.1 (3500.0)
Macaque							
prefro. (young)	0.106±0.660[m]	1.93 (2.92)	2.83 (2.83)	-43.1 (3.04)	65.9 (73.6)	99.9 (0.9)	- (97.4)
prefro. (old)	0.110±0.527[m]	2.19 (2.93)	2.84 (2.84)	-21.7 (3.05)	74.9 (62.2)	99.9 (1.7)	- (95.6)
prefro.	0.102±0.075[n]	2.72 (2.91)	2.56 (2.82)	2.89 (3.03)	93.3 (25.3)	90.8 (47.3)	95.5 (85.0)
visual (young)	0.062±0.173[m]	2.50 (2.79)	2.69 (2.70)	-2.79 (2.91)	89.7 (35.8)	99.8 (4.9)	- (87.2)
visual (old)	0.070±0.226[m]	2.43 (2.82)	2.73 (2.73)	-5.83 (2.94)	86.0 (42.8)	99.9 (3.7)	- (90.4)
visual	0.066±0.141[n]	2.68 (2.81)	2.70 (2.71)	0.61 (2.92)	95.5 (23.3)	99.6 (8.1)	20.7 (78.1)
cingul.	0.07±0.15[o]	2.69 (2.82)	2.72 (2.73)	0.59 (2.94)	95.4 (23.5)	99.6 (8.1)	20.2 (78.2)
pariet. (young)	0.196±0.353[p]	3.02 (3.07)	2.96 (2.98)	2.03 (3.19)	98.2 (16.1)	99.3 (11.1)	63.7 (69.2)
pariet. (old)	0.255±1.528[p]	2.18 (3.14)	3.04 (3.05)	-38.9 (3.25)	69.5 (80.5)	99.9 (0.9)	- (97.2)
Human							
cingul. (40y api.)	0.301±0.245[r]	3.04 (3.18)	2.88 (3.09)	3.23 (3.29)	95.8 (22.7)	93.4 (40.8)	99.0 (50.9)
cingul. (40y bas.)	0.284±0.232[r]	3.03 (3.16)	2.87 (3.07)	3.22 (3.28)	95.8 (22.4)	93.4 (40.6)	98.1 (49.9)
cingul. (85y api.)	0.310±0.244[r]	3.03 (3.18)	2.87 (3.09)	3.21 (3.30)	95.2 (24.2)	92.8 (42.9)	97.4 (61.4)
cingul. (85y bas.)	0.325±0.277[r]	3.09 (3.20)	2.92 (3.11)	3.27 (3.31)	96.6 (20.9)	94.1 (38.1)	98.8 (37.7)

[a] Ishii et al (2018); [b] Tamada et al (2020); [c] Kashiwagi et al (2019); [d] Schikorski and Stevens (1999); [e] Rodriguez-Moreno et al (2018); [f] Arellano et al (2007); [g] Parajuli et al (2020); [h] Harris and Stevens (1989); [i] Borczyk (2019); [j] Radley et al (2008); [k] Cheetham et al (2014); [l] Harris and Stevens (1988); [m] Young et al (2014); [n] Medalla and Luebke (2015); [o] Medalla and Barbas (2009); [p] Motley et al (2018) (and J.H. Morrison, private comm.); [r] Benavides-Piccione (2013; and private comm.).

Table 5: Information contained in the distribution of spine/PSD areas.

Species/ region (cond.)	Spine area mean±SD (μm^2)	Entropy (bits)			Efficiency % (Deviation %)		
		$H_{ln}(H_{ln,m})$	$H_{ll}(H_{ll,m})$	$H_g(H_{g,m})$	$\eta_{ln}(D_{ln})$	$\eta_{ll}(D_{ll})$	$\eta_g(D_g)$
Mouse							
hippo.	0.047±0.038[a]	2.25 (2.39)	2.09 (2.30)	2.44 (2.51)	94.3 (21.1)	90.9 (41.2)	97.2 (53.0)
hippo. (PSD)	0.043±0.031[b]	2.15 (2.36)	2.00 (2.27)	2.32 (2.48)	91.1 (25.6)	88.1 (48.7)	93.9 (92.4)
piriform 1a (PSD)	0.096±0.105[b]	2.63 (2.65)	2.46 (2.56)	2.75 (2.77)	99.3 (8.4)	96.2 (25.8)	99.4 (16.4)
piriform 1b (PSD)	0.100±0.087[b]	2.56 (2.66)	2.40 (2.57)	2.75 (2.78)	96.3 (18.5)	93.2 (36.9)	98.9 (32.1)
visual (PSD)	0.08±0.06[c]	2.40 (2.58)	2.24 (2.49)	2.58 (2.70)	92.9 (24.3)	90.1 (46.0)	95.5 (77.8)
visual	0.31±0.35[d]	3.06 (3.07)	2.89 (2.98)	3.16 (3.19)	99.6 (7.6)	97.0 (24.5)	99.0 (21.6)
temporal	0.37±0.36[d]	3.08 (3.13)	2.91 (3.04)	3.25 (3.25)	98.3 (15.2)	95.7 (31.1)	99.9 (5.6)
somatosen.	0.87±0.47[e]	3.00 (3.44)	2.88 (3.35)	3.13 (3.56)	87.0 (53.9)	85.9 (73.1)	87.8 (242.6)
Rat							
hippo.	0.83±0.63[f]	3.25 (3.43)	3.09 (3.34)	3.43 (3.54)	94.9 (36.3)	92.8 (45.2)	96.8 (73.6)
hippo. (PSD)	0.069±0.080[f]	2.52 (2.53)	2.36 (2.44)	2.60 (2.65)	99.7 (5.7)	96.7 (23.5)	98.2 (25.6)
hippo. (PSD; ctr)	0.068±0.073[g]	2.50 (2.52)	2.33 (2.43)	2.63 (2.64)	99.1 (9.2)	95.8 (26.6)	99.6 (13.2)
hippo. (PSD; LTP)	0.109±0.179[g]	2.67 (2.69)	2.58 (2.60)	2.07 (2.81)	98.9 (10.9)	98.9 (13.0)	73.5 (62.9)
prefro. (ctr)	1.1±1.8[h]	3.50 (3.53)	3.41 (3.44)	2.91 (3.64)	99.2 (28.2)	99.2 (13.1)	79.9 (62.7)
prefro. (stress)	1.0±1.9[h]	3.42 (3.49)	3.38 (3.40)	2.15 (3.61)	97.8 (40.9)	99.5 (10.0)	59.6 (72.3)
forebr. (PSD; E19d)	0.195±0.118[i]	2.56 (2.90)	2.43 (2.81)	2.71 (3.02)	88.2 (33.2)	86.3 (62.5)	89.6 (173.1)
forebr. (PSD; P2d)	0.313±0.216[i]	2.83 (3.07)	2.69 (2.98)	3.00 (3.19)	92.2 (29.6)	90.1 (51.9)	94.1 (110.0)
forebr. (PSD; P21d)	0.259±0.214[i]	2.88 (3.01)	2.72 (2.92)	3.07 (3.12)	95.8 (21.7)	93.2 (39.9)	98.2 (46.5)
forebr. (PSD; P60d)	0.251±0.153[i]	2.66 (2.99)	2.52 (2.90)	2.81 (3.11)	88.7 (33.5)	86.9 (61.9)	90.2 (169.1)
cerebel.	1.12±0.18[j]	1.59 (3.53)	1.56 (3.44)	1.61 (3.65)	45.0 (107.1)	45.2 (336.7)	44.0 (3771.6)
cerebel. (PSD)	0.15±0.08[j]	2.35 (2.81)	2.23 (2.72)	2.48 (2.93)	83.7 (37.0)	82.2 (74.4)	84.6 (251.6)
Cat							
visual	0.099±0.046[k]	2.06 (2.66)	1.96 (2.57)	2.17 (2.78)	77.5 (40.7)	76.4 (89.7)	78.0 (363.2)
Macaque							
prefro. (PSD)	0.12±0.13[l]	2.71 (2.73)	2.54 (2.64)	2.83 (2.85)	99.2 (8.9)	96.3 (26.2)	99.5 (14.8)
visual (PSD)	0.08±0.18[l]	2.43 (2.58)	2.48 (2.49)	-0.10 (2.70)	94.0 (26.0)	99.6 (7.4)	- (80.2)
cingul. (PSD)	0.13±0.31[m]	2.57 (2.76)	2.66 (2.67)	-0.55 (2.87)	93.2 (29.5)	99.7 (6.6)	- (82.4)
Human							
neoctx (PSD)	0.205±0.121[n]	2.56 (2.92)	2.43 (2.83)	2.70 (3.04)	87.5 (34.2)	85.8 (64.7)	88.9 (187.0)
temporal	0.59±0.53[d]	3.22 (3.30)	3.05 (3.21)	3.40 (3.42)	97.5 (21.5)	95.0 (35.2)	99.4 (23.9)

[a] Santuy et al (2020); [b] Schikorski and Stevens (1999); [c] Arellano et al (2007); [d] Benavides-Piccione et al (2002); [e] Rodriguez-Moreno et al (2018); [f] Harris and Stevens (1989); [g] Borczyk et al (2019); [h] Radley et al (2008); [i] Swulius et al (2010); [j] Harris and Stevens (1988); [k] da Costa (2013); [l] Medalla and Luebke (2015); [m] Medalla and Barbas (2009); [n] Tang et al (2001).

Table 6: Information contained in the distribution of spine length.

Species/ region (cond.)	Spine length mean±SD (μm)	Entropy (bits)			Efficiency % (Deviation %)		
		$H_{ln}(H_{ln,m})$	$H_{ll}(H_{ll,m})$	$H_g(H_{g,m})$	$\eta_{ln}(D_{ln})$	$\eta_{ll}(D_{ll})$	$\eta_g(D_g)$
Mouse							
neoctx	0.836±0.569[a]	4.53 (4.78)	4.38 (4.68)	4.69 (4.89)	94.8 (42.1)	93.5 (52.9)	95.9 (115.9)
hippo.	0.95±0.42[b]	4.22 (4.87)	4.12 (4.78)	4.31 (4.98)	86.6 (66.7)	86.3 (95.9)	86.5 (411.6)
hippo.	0.83±0.27[c]	3.76 (4.77)	3.70 (4.68)	3.82 (4.89)	78.9 (66.9)	79.0 (143.0)	78.1 (845.0)
hippo. (neck)	0.42±0.24[c]	3.88 (4.28)	3.76 (4.19)	4.02 (4.40)	90.8 (38.0)	89.8 (67.7)	91.5 (206.2)
visual (neck)	0.66±0.37[d]	4.19 (4.60)	4.07 (4.51)	4.33 (4.72)	91.1 (43.9)	90.2 (69.5)	91.7 (218.2)
striat. (wt, 1m)	1.00±0.52[e]	4.42 (4.90)	4.31 (4.81)	4.54 (5.02)	90.2 (64.8)	89.5 (77.1)	90.5 (269.8)
striat. (mut, 1m)	1.14±0.57[e]	4.48 (5.00)	4.37 (4.91)	4.59 (5.12)	89.6 (83.3)	89.0 (81.3)	89.8 (299.9)
striat. (wt, 3m)	1.13±0.57[e]	4.48 (4.99)	4.37 (4.90)	4.60 (5.11)	89.7 (81.3)	89.2 (80.3)	89.9 (293.0)
striat. (mut, 3m)	1.11±0.57[e]	4.48 (4.98)	4.37 (4.89)	4.60 (5.10)	90.1 (77.6)	89.4 (78.4)	90.3 (279.2)
striat. (wt, 6m)	0.98±0.56[e]	4.50 (4.89)	4.37 (4.80)	4.63 (5.01)	91.9 (58.9)	91.1 (67.7)	92.5 (206.3)
striat. (mut, 6m)	1.21±0.63[e]	4.56 (5.04)	4.45 (4.95)	4.68 (5.16)	90.5 (94.1)	89.8 (76.9)	90.8 (268.9)
striat. (wt, 22m)	1.03±0.54[e]	4.45 (4.93)	4.34 (4.84)	4.57 (5.04)	90.4 (67.3)	89.7 (76.2)	90.7 (263.8)
striat. (mut, 22m)	1.11±0.57[e]	4.48 (4.98)	4.37 (4.89)	4.60 (5.10)	90.1 (77.6)	89.4 (78.4)	90.3 (279.2)
Rat							
hippo.	0.95±0.42[f]	4.22 (4.87)	4.12 (4.78)	4.31 (4.98)	86.6 (66.7)	86.3 (95.9)	86.5 (411.6)
prefro. (ctr)	0.55±1.80[g]	4.07 (4.47)	4.38 (4.38)	-4.52 (4.59)	91.0 (61.9)	99.9 (3.6)	- (90.7)
prefro. (stress)	0.55±0.95[g]	4.43 (4.47)	4.36 (4.38)	3.64 (4.59)	99.1 (17.5)	99.5 (11.9)	79.2 (66.5)
cerebel.	1.22±0.30[h]	3.68 (5.05)	3.63 (4.96)	3.72 (5.17)	72.9 (122.8)	73.3 (202.8)	72.0 (1553.8)
Cat							
visual	1.03±0.74[i]	4.71 (4.93)	4.56 (4.84)	4.89 (5.04)	95.7 (50.5)	94.4 (49.0)	96.9 (93.7)
Macaque							
prefro. (young)	1.481±2.968[j]	5.09 (5.19)	5.08 (5.10)	3.49 (5.31)	98.1 (202.8)	99.7 (9.1)	65.8 (75.1)
prefro. (old)	1.336±3.516[j]	4.87 (5.11)	5.02 (5.02)	0.49 (5.23)	95.2 (182.6)	99.9 (5.5)	9.3 (85.6)
visual (young)	1.388±3.179[j]	4.98 (5.14)	5.04 (5.05)	2.28 (5.26)	96.8 (172.7)	99.8 (7.1)	43.3 (80.9)
visual (old)	1.462±2.883[j]	5.09 (5.18)	5.07 (5.09)	3.60 (5.30)	98.3 (173.6)	99.7 (9.4)	67.9 (74.3)
visual (young)	1.10±1.32[k]	4.97 (4.97)	4.81 (4.88)	5.02 (5.09)	99.9 (10.2)	98.5 (22.2)	98.6 (30.6)
visual (old)	1.20±2.43[k]	4.93 (5.04)	4.93 (4.95)	3.27 (5.15)	98.0 (72.7)	99.7 (9.0)	63.4 (75.6)
Human							
cingul. (40y api.)	1.780±0.740[l]	4.60 (5.32)	4.51 (5.23)	4.69 (5.44)	86.5 (390.2)	86.3 (104.0)	86.2 (478.6)
cingul. (40y bas.)	1.417±0.710[l]	4.64 (5.16)	4.53 (5.07)	4.75 (5.27)	89.9 (184.9)	89.4 (81.0)	90.1 (298.3)
cingul. (85y api.)	1.709±0.769[l]	4.66 (5.29)	4.56 (5.20)	4.76 (5.41)	88.1 (804.1)	87.8 (93.7)	88.0 (393.9)
cingul. (85y bas.)	1.483±0.715[l]	4.63 (5.19)	4.53 (5.10)	4.74 (5.31)	89.2 (266.8)	88.8 (85.4)	89.3 (330.2)
prefro. (non-alzh.)	1.219±2.335[m]	4.97 (5.05)	4.94 (4.96)	3.66 (5.16)	98.4 (65.6)	99.7 (9.9)	70.8 (72.7)
prefro. (alzh.)	1.360±2.980[m]	4.99 (5.13)	5.02 (5.04)	2.70 (5.24)	97.3 (141.1)	99.8 (7.7)	51.5 (79.2)
hippo. (5m)	1.347±0.697[n]	4.63 (5.12)	4.52 (5.03)	4.75 (5.24)	90.5 (138.3)	89.9 (77.6)	90.8 (273.5)
hippo. (2y)	1.359±0.662[n]	4.58 (5.13)	4.47 (5.04)	4.69 (5.24)	89.3 (148.9)	88.8 (84.2)	89.4 (321.4)
hippo. (23y)	1.376±0.650[n]	4.56 (5.13)	4.45 (5.04)	4.66 (5.25)	88.7 (160.9)	88.3 (87.8)	88.7 (348.1)
hippo. (27y)	1.390±0.669[n]	4.58 (5.14)	4.48 (5.05)	4.69 (5.26)	89.1 (168.4)	88.6 (85.6)	89.2 (331.7)
hippo. (38y)	1.383±0.663[n]	4.57 (5.14)	4.47 (5.05)	4.68 (5.26)	89.0 (164.1)	88.6 (86.0)	89.1 (335.1)
hippo. (45y)	1.358±0.659[n]	4.57 (5.13)	4.47 (5.03)	4.68 (5.24)	89.2 (148.7)	88.8 (84.6)	89.3 (324.6)
hippo. (57y)	1.346±0.651[n]	4.56 (5.12)	4.46 (5.03)	4.67 (5.24)	89.2 (142.9)	88.7 (85.0)	89.3 (327.5)
hippo. (58y)	1.331±0.649[n]	4.56 (5.11)	4.46 (5.02)	4.68 (5.23)	89.3 (135.4)	88.8 (84.1)	89.4 (320.6)
hippo. (68y)	1.342±0.655[n]	4.57 (5.12)	4.47 (5.03)	4.68 (5.23)	89.3 (140.3)	88.8 (84.0)	89.5 (319.8)
hippo. (70y)	1.348±0.654[n]	4.57 (5.12)	4.46 (5.03)	4.68 (5.24)	89.2 (143.6)	88.7 (84.7)	89.3 (324.8)
hippo. (71y)	1.341±0.649[n]	4.56 (5.12)	4.46 (5.03)	4.67 (5.23)	89.2 (140.4)	88.7 (84.9)	89.3 (326.9)

[a] Tamada et al (2020); [b] Kashiwagi et al (2019); [c] Scheuss and Bonhoeffer (2014); [d] Arellano et al (2007); [e] Parajuli et al (2020); [f] Harris and Stevens (1989); [g] Radley et al (2008); [h] Harris and Stevens (1988); [i] da Costa (2013); [j] Young et al (2014); [k] Luebke et al (2015); [l] Benavides-Piccione (2013; and private comm.); [m] Boros et al (2019; and J.H. Herskowitz, private comm.) ; [n] Das et al (2019; and S.C. Das, private comm.).

Table 7: Information contained in the distribution of spine head/PSD diameter.

Species/ region (cond.)	Spine diameter mean±SD (μm)	Entropy (bits)			Efficiency % (Deviation %)		
		$H_{ln}(H_{ln,m})$	$H_{ll}(H_{ll,m})$	$H_g(H_{g,m})$	$\eta_{ln}(D_{ln})$	$\eta_{ll}(D_{ll})$	$\eta_g(D_g)$
Mouse							
hippo.	0.66±0.11[a]	2.71 (4.60)	2.68 (4.51)	2.73 (4.72)	58.9 (69.9)	59.3 (322.7)	57.8 (3500.0)
visual	0.42±0.13[b]	3.21 (4.28)	3.15 (4.19)	3.26 (4.40)	75.0 (54.6)	75.1 (152.3)	74.2 (943.8)
Rat							
hippo.	0.53±0.28[c]	3.98 (4.45)	3.86 (4.36)	4.10 (4.56)	89.5 (42.7)	88.7 (75.4)	89.9 (258.3)
cerebel.	0.57±0.05[d]	1.70 (4.50)	1.73 (4.41)	1.72 (4.62)	37.8 (72.5)	39.3 (637.6)	37.3 (12895.9)
Rabbit							
hippo. (E21d)	0.33±0.09[e]	2.87 (4.10)	2.82 (4.01)	2.92 (4.22)	70.0 (55.6)	70.3 (178.5)	69.1 (1244.4)
hippo. (E24d)	0.49±0.13[e]	3.12 (4.39)	3.07 (4.30)	3.16 (4.51)	71.1 (58.9)	71.4 (184.7)	70.2 (1320.7)
hippo. (E28d)	0.55±0.15[e]	3.24 (4.47)	3.19 (4.38)	3.28 (4.59)	72.5 (59.8)	72.8 (178.5)	71.5 (1244.4)
hippo. (P2d)	0.48±0.11[e]	2.92 (4.38)	2.87 (4.28)	2.95 (4.49)	66.6 (61.1)	67.0 (221.0)	65.6 (1804.1)
hippo. (P8d)	0.43±0.10[e]	2.86 (4.30)	2.81 (4.21)	2.89 (4.41)	66.5 (59.9)	66.8 (217.1)	65.4 (1749.0)
hippo. (P14d)	0.42±0.11[e]	3.00 (4.28)	2.94 (4.19)	3.03 (4.40)	70.0 (57.8)	70.3 (187.7)	69.0 (1357.9)
hippo. (adult)	0.41±0.10[e]	2.88 (4.26)	2.84 (4.17)	2.92 (4.38)	67.7 (58.8)	68.0 (204.8)	66.7 (1581.0)
Echidna							
visual	0.323±0.120[f]	3.24 (4.09)	3.16 (4.00)	3.31 (4.21)	79.3 (49.1)	79.1 (120.5)	78.8 (624.5)
Cat							
auditory (hear)	0.672±0.213[g]	3.58 (4.62)	3.51 (4.53)	3.63 (4.74)	77.5 (60.4)	77.6 (147.8)	76.7 (895.4)
auditory (deaf)	0.707±0.233[g]	3.66 (4.65)	3.60 (4.56)	3.72 (4.77)	78.7 (60.9)	78.8 (140.6)	78.0 (820.7)
Macaque							
prefro. (young)	0.403±0.660[h]	4.22 (4.25)	4.13 (4.16)	3.63 (4.37)	99.3 (12.6)	99.3 (13.1)	83.2 (62.7)
prefro. (old)	0.420±0.440[h]	4.25 (4.28)	4.08 (4.19)	4.39 (4.40)	99.3 (11.9)	97.4 (27.7)	99.9 (8.9)
visual (young)	0.376±0.289[h]	4.03 (4.20)	3.87 (4.11)	4.21 (4.32)	96.0 (25.9)	94.3 (44.4)	97.6 (69.3)
visual (old)	0.383±0.226[h]	3.85 (4.21)	3.72 (4.12)	3.99 (4.33)	91.3 (36.3)	90.2 (64.7)	92.2 (187.2)
visual (young)	0.43±0.53[i]	4.29 (4.30)	4.14 (4.21)	4.31 (4.41)	99.9 (3.4)	98.4 (21.3)	97.8 (34.2)
visual (old)	0.46±0.30[i]	4.06 (4.34)	3.92 (4.25)	4.22 (4.46)	93.5 (33.7)	92.2 (56.2)	94.6 (135.1)
Dolphin							
visual	0.352±0.140[j]	3.38 (4.15)	3.30 (4.06)	3.46 (4.27)	81.5 (47.8)	81.3 (110.3)	81.1 (532.2)
Human							
neoctx	0.466±0.121[k]	3.06 (4.35)	3.01 (4.26)	3.10 (4.47)	70.3 (58.8)	70.6 (189.7)	69.3 (1383.2)
temporal	0.30±0.09[l]	2.93 (4.04)	2.87 (3.95)	2.98 (4.15)	72.5 (53.4)	72.7 (158.4)	71.7 (1011.1)
prefro. (non-alzh.)	0.573±0.803[m]	4.50 (4.50)	4.37 (4.41)	4.33 (4.62)	99.9 (4.2)	99.0 (17.2)	93.7 (49.1)
prefro. (alzh.)	0.562±1.141[m]	4.39 (4.49)	4.38 (4.40)	2.70 (4.61)	97.7 (28.6)	99.7 (8.9)	58.6 (75.7)
hippo. (5m)	0.375±0.166[n]	3.55 (4.20)	3.45 (4.11)	3.64 (4.31)	84.5 (45.3)	84.1 (95.7)	84.5 (410.3)
hippo. (2y)	0.380±0.175[n]	3.60 (4.21)	3.50 (4.12)	3.70 (4.32)	85.6 (44.2)	85.1 (90.8)	85.6 (371.5)
hippo. (23y)	0.386±0.175[n]	3.59 (4.22)	3.50 (4.13)	3.69 (4.34)	85.2 (44.8)	84.7 (92.7)	85.2 (386.5)
hippo. (27y)	0.391±0.181[n]	3.63 (4.23)	3.53 (4.14)	3.73 (4.34)	85.8 (44.2)	85.3 (90.2)	85.8 (366.7)
hippo. (38y)	0.389±0.184[n]	3.65 (4.22)	3.54 (4.13)	3.75 (4.34)	86.3 (43.6)	85.7 (87.6)	86.4 (346.9)
hippo. (45y)	0.391±0.181[n]	3.63 (4.23)	3.53 (4.14)	3.73 (4.34)	85.8 (44.2)	85.3 (90.2)	85.8 (366.7)
hippo. (57y)	0.391±0.179[n]	3.61 (4.23)	3.52 (4.14)	3.72 (4.34)	85.5 (44.6)	85.0 (91.5)	85.5 (377.1)
hippo. (58y)	0.385±0.177[n]	3.61 (4.22)	3.51 (4.13)	3.71 (4.33)	85.6 (44.3)	85.1 (91.0)	85.6 (373.1)
hippo. (68y)	0.387±0.181[n]	3.63 (4.22)	3.53 (4.13)	3.73 (4.34)	86.0 (43.9)	85.5 (88.9)	86.1 (357.2)
hippo. (70y)	0.387±0.183[n]	3.64 (4.22)	3.54 (4.13)	3.75 (4.34)	86.3 (43.5)	85.7 (87.6)	86.4 (347.2)
hippo. (71y)	0.390±0.185[n]	3.65 (4.23)	3.55 (4.13)	3.76 (4.34)	86.4 (43.5)	85.8 (87.3)	86.5 (344.4)

[a] Scheuss and Bonhoeffer (2014); [b] Konur et al (2003); [c] Harris and Stevens (1989); [d] Harris and Stevens (1988); [e] Schwartzkroin (1982); [f] Hassiotis et al (2003); [g] Clemo et al (2016); [h] Young et al (2014); [i] Luebke et al (2015); [j] Glezer and Morgane (1990); [k] Tang et al (2001); [l] Alonso-Nanclares et al (2008); [m] Boros et al (2019; and private comm.) [n] Das et al (2019; and private comm.).

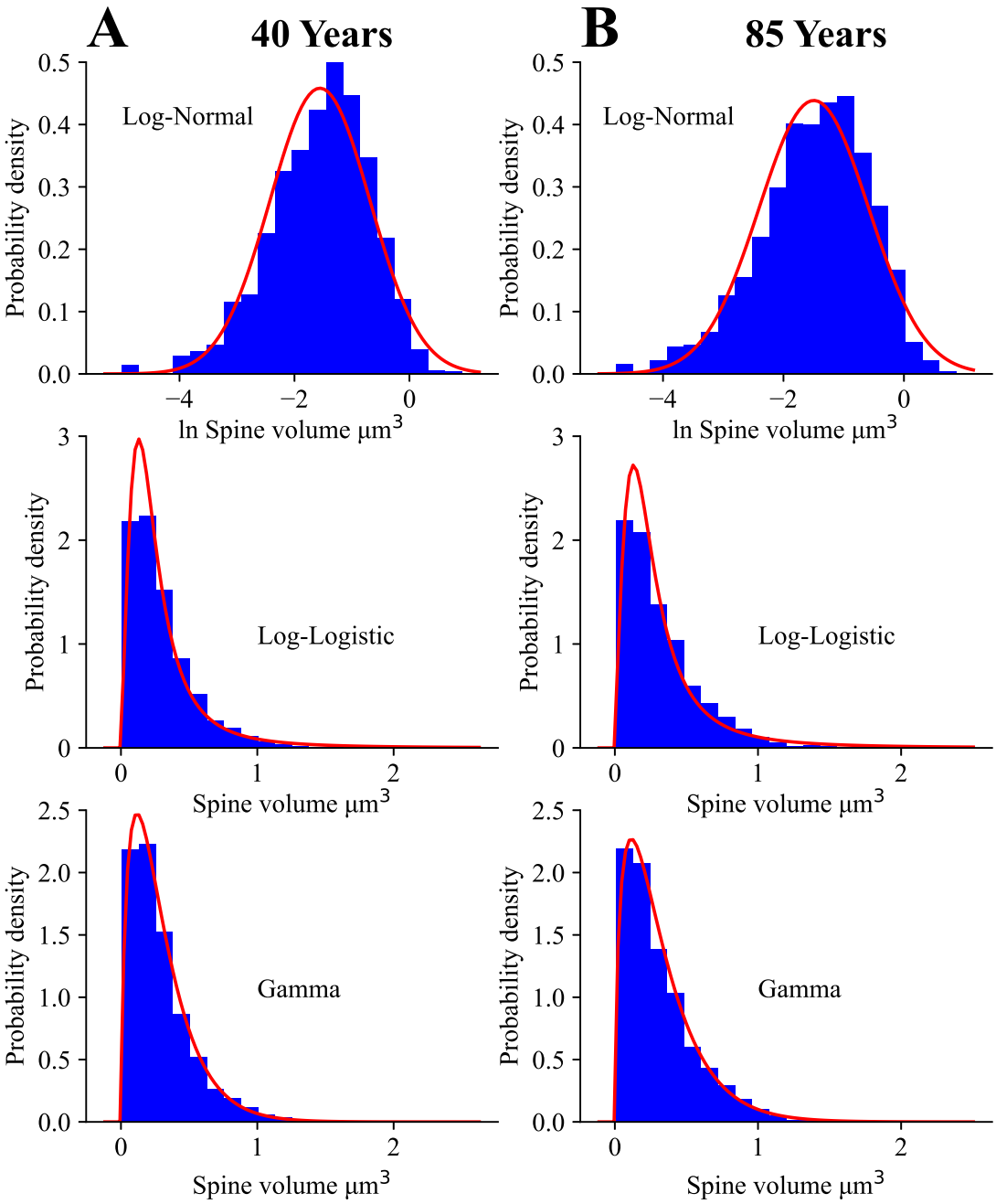
A
40 Years
B
85 Years
Log-Normal
Log-Normal
Probability density
ln Spine volume μm3
Log-Logistic
Log-Logistic
Spine volume μm3
Gamma
Gamma

Figure 1

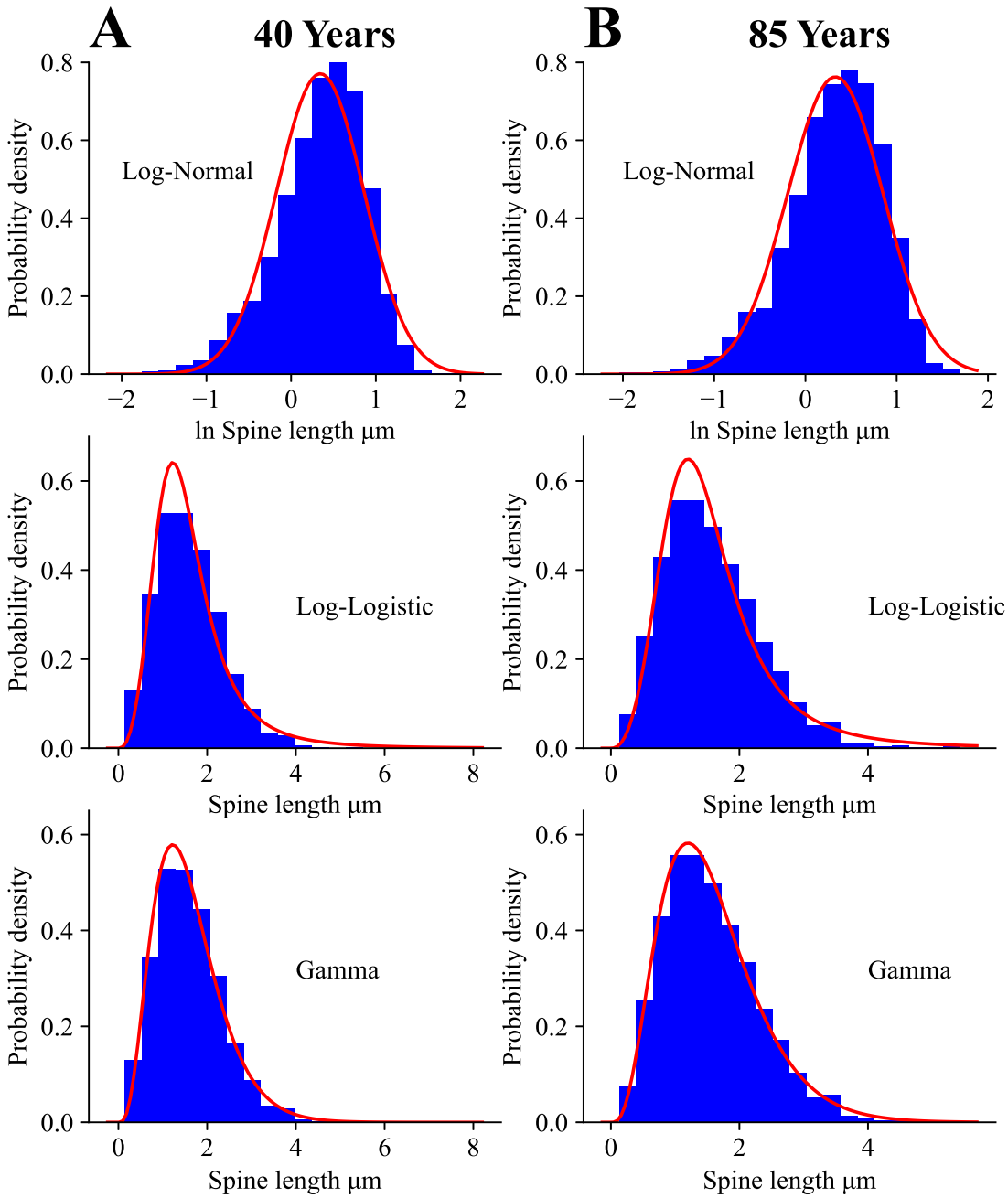
A
40 Years
B
85 Years
Log-Normal
Log-Logistic
Gamma
Probability density
ln Spine length µm
Spine length µm

Figure 2

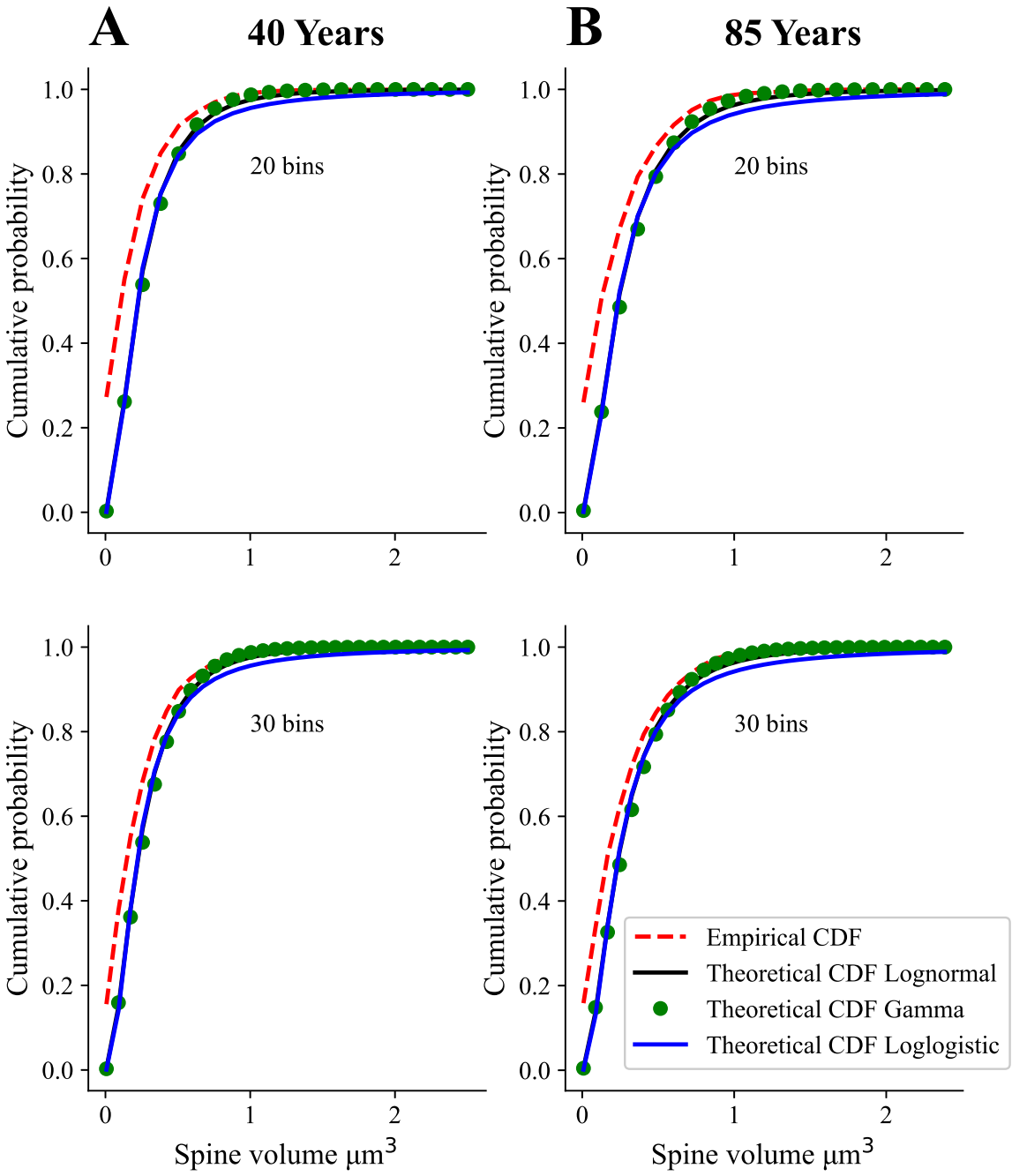
A
40 Years
B
85 Years
20 bins
30 bins
Cumulative probability
Spine volume μm3
Empirical CDF
Theoretical CDF Lognormal
Theoretical CDF Gamma
Theoretical CDF Loglogistic

Figure 3

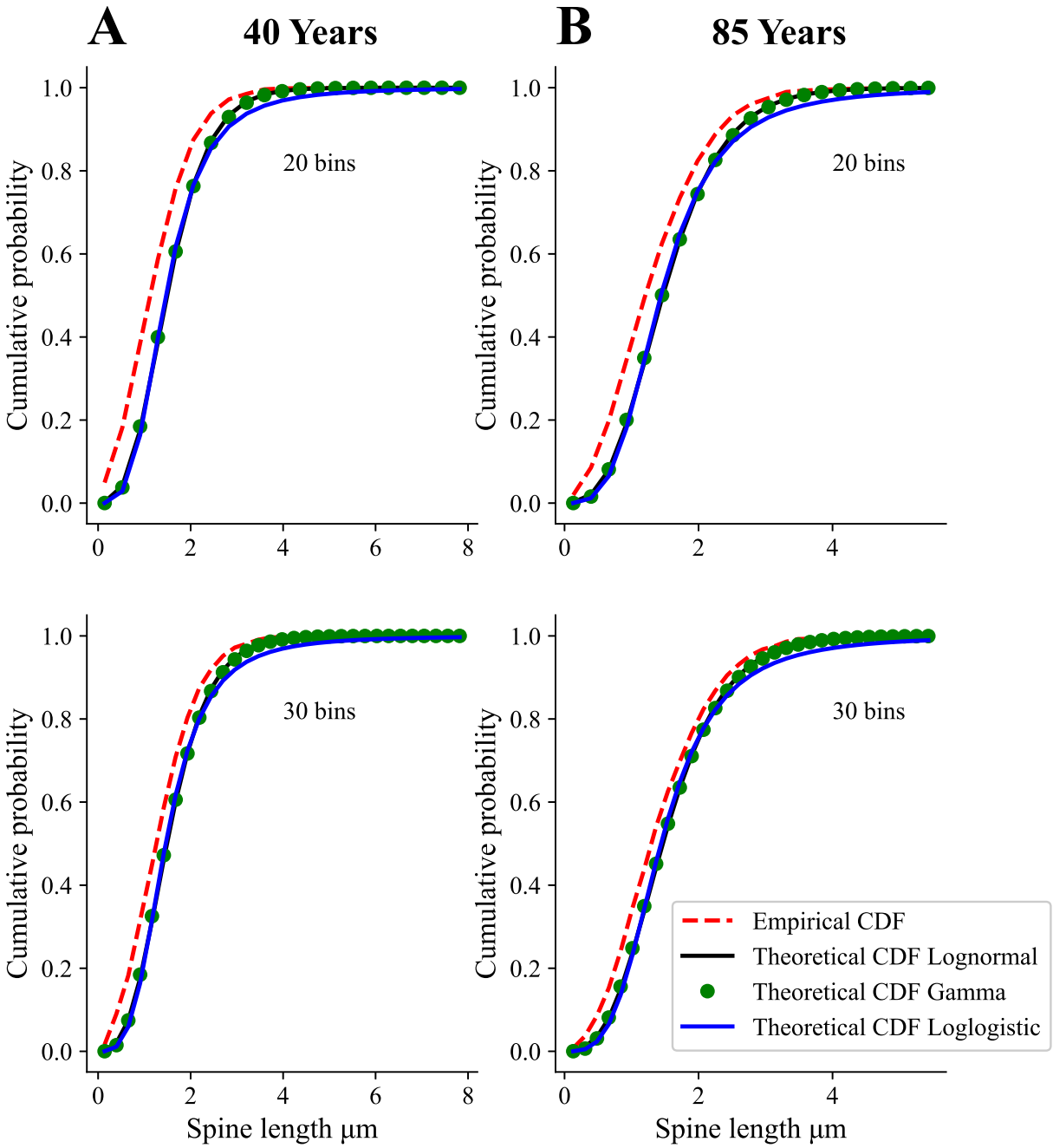
A
40 Years
B
85 Years
20 bins
30 bins
Cumulative probability
Spine length µm
Empirical CDF
Theoretical CDF Lognormal
Theoretical CDF Gamma
Theoretical CDF Loglogistic

Figure 4

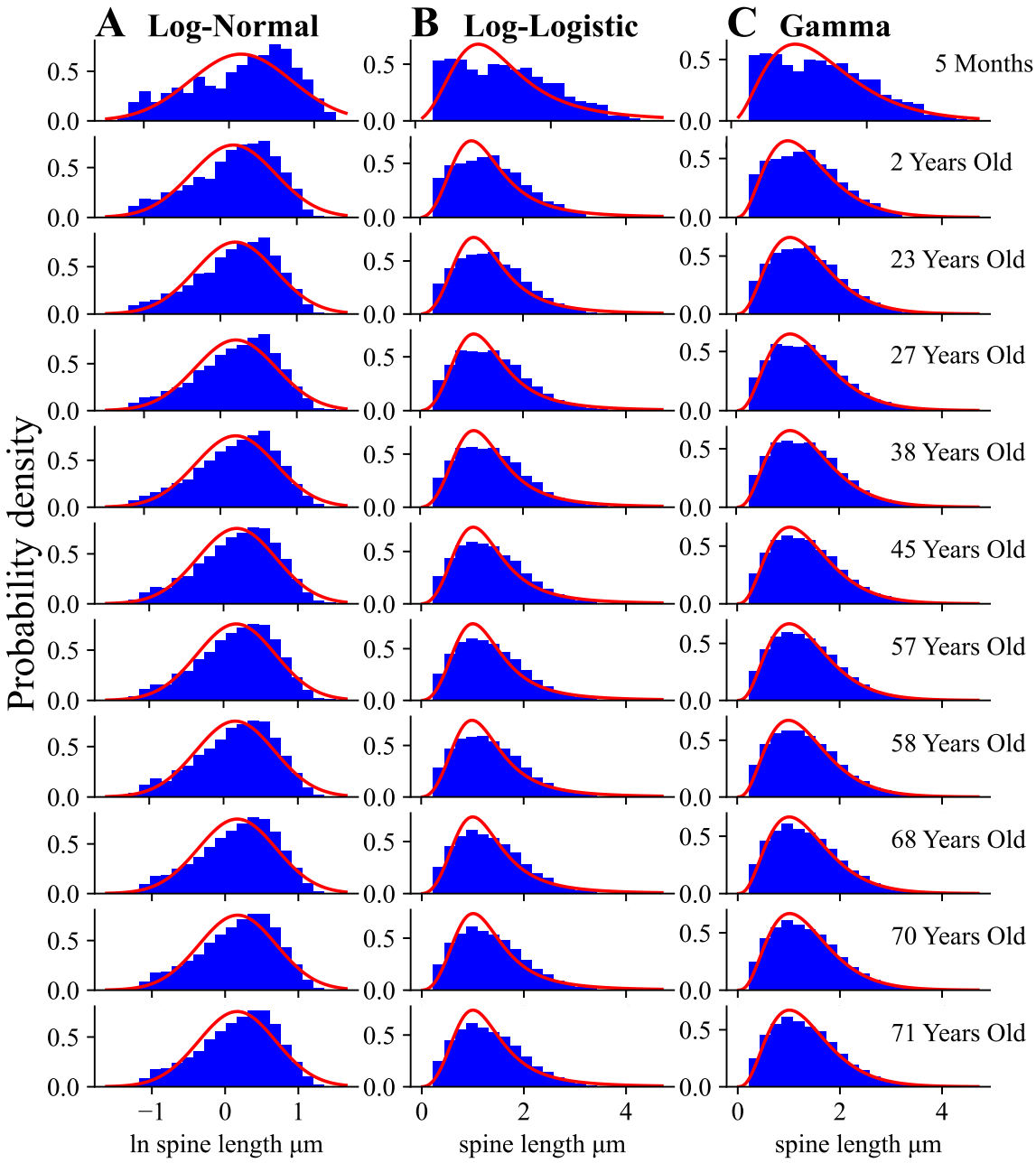
A Log-Normal
B Log-Logistic
C Gamma
5 Months
2 Years Old
23 Years Old
27 Years Old
38 Years Old
45 Years Old
57 Years Old
58 Years Old
68 Years Old
70 Years Old
71 Years Old
Probability density
0.0
0.5
−1
0
1
2
4
ln spine length µm
spine length µm

Figure 5

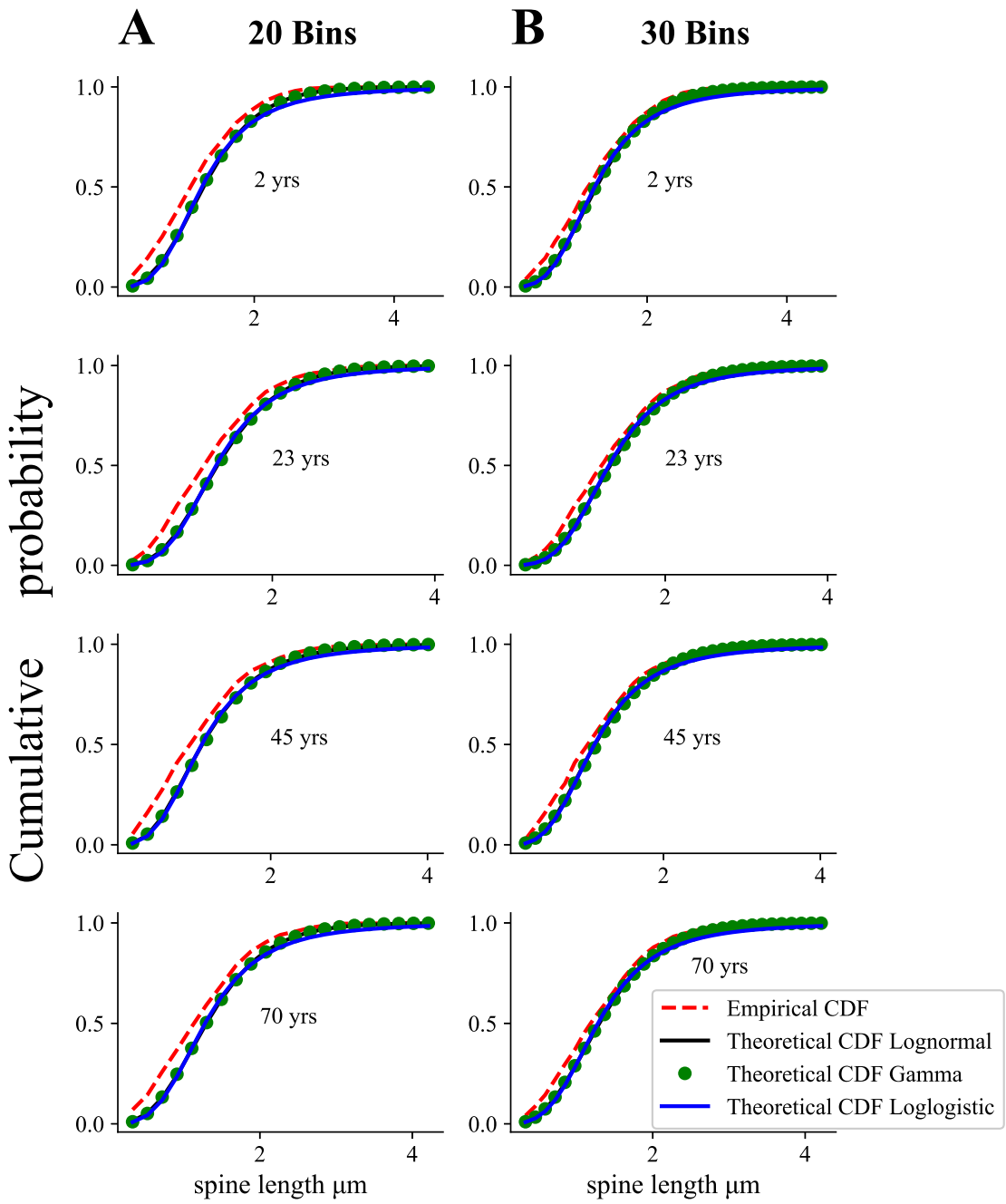
A
20 Bins
B
30 Bins
Cumulative probability
2 yrs
23 yrs
45 yrs
70 yrs
1.0
0.5
0.0
2
4
spine length µm
Empirical CDF
Theoretical CDF Lognormal
Theoretical CDF Gamma
Theoretical CDF Loglogistic

Figure 6

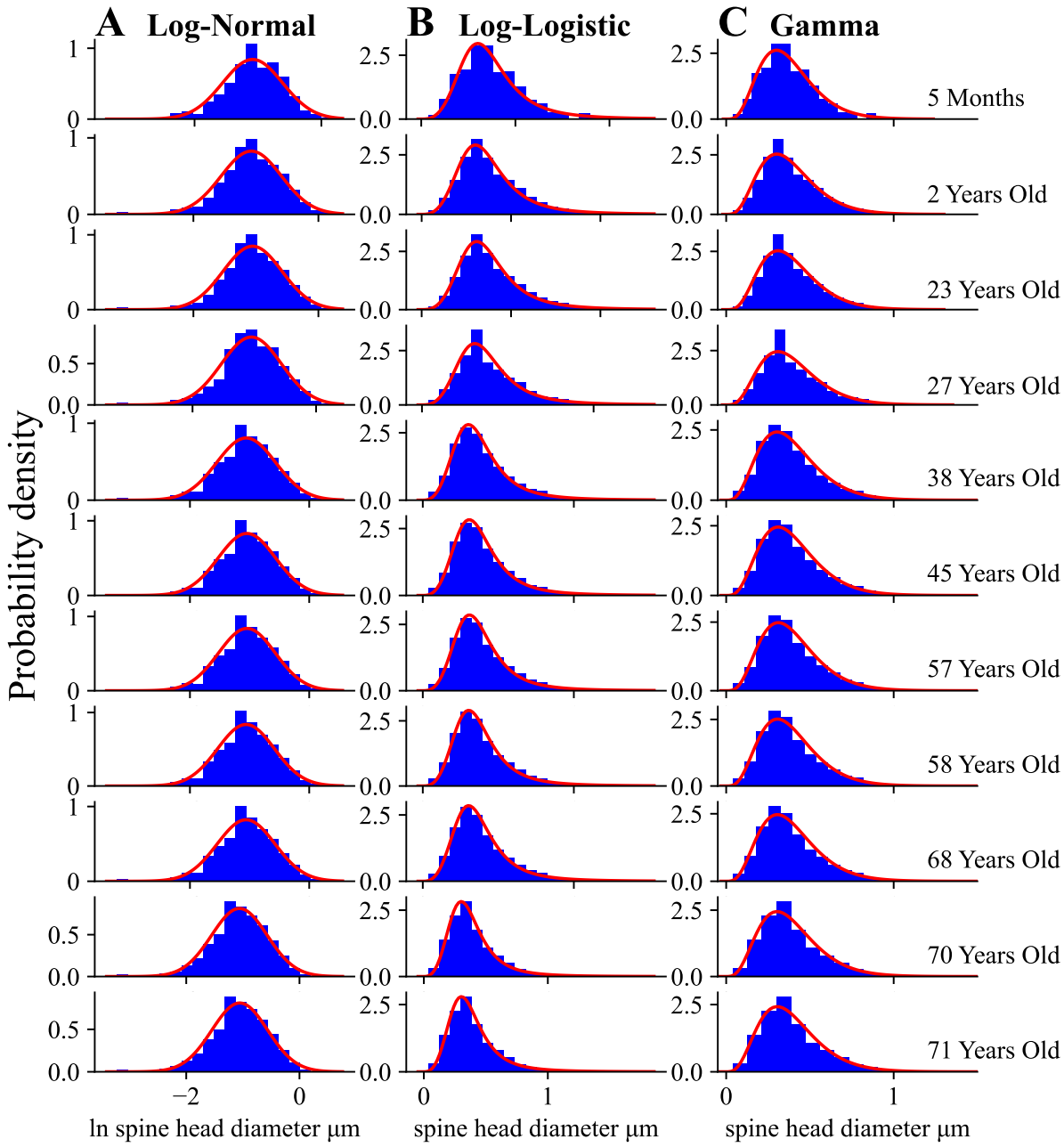
A Log-Normal
B Log-Logistic
C Gamma
Probability density
5 Months
2 Years Old
23 Years Old
27 Years Old
38 Years Old
45 Years Old
57 Years Old
58 Years Old
68 Years Old
70 Years Old
71 Years Old
ln spine head diameter µm
spine head diameter µm
spine head diameter µm

Figure 7

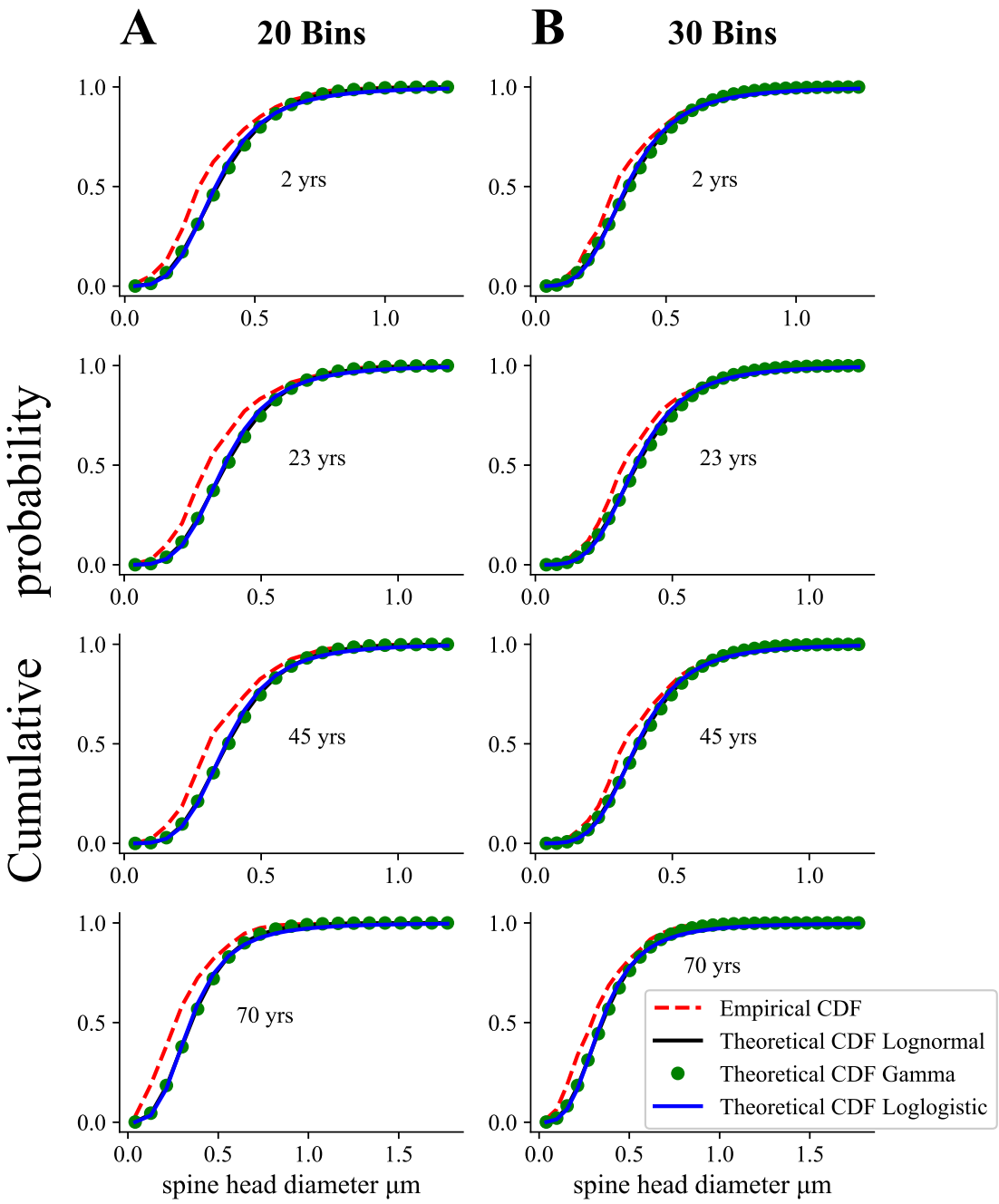
A
20 Bins
B
30 Bins
Cumulative probability
2 yrs
23 yrs
45 yrs
70 yrs
Empirical CDF
Theoretical CDF Lognormal
Theoretical CDF Gamma
Theoretical CDF Loglogistic
spine head diameter µm

Figure 8

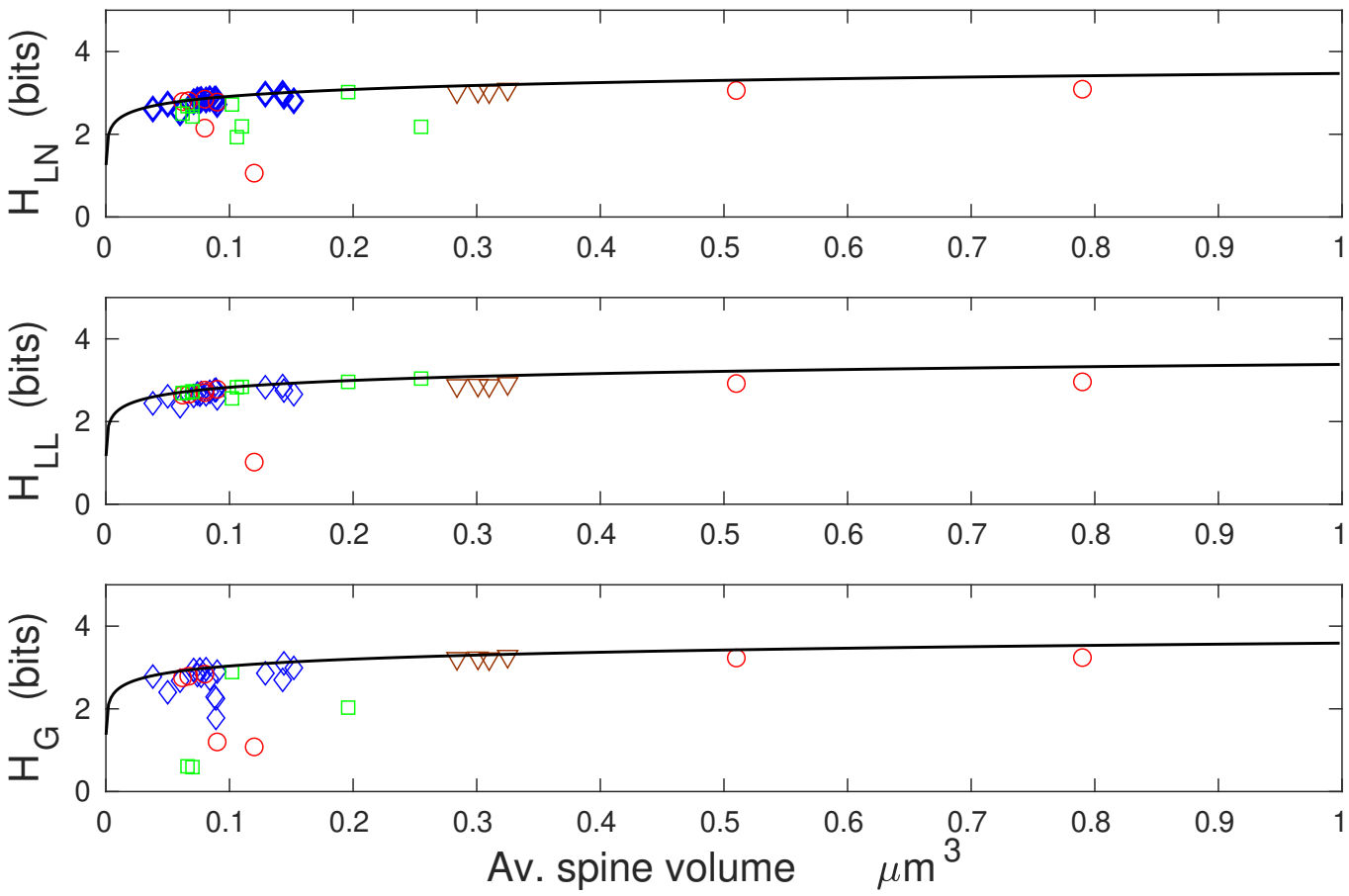
H_LN (bits)
H_LL (bits)
H_G (bits)
0
2
4
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Av. spine volume μm^3

Figure 9

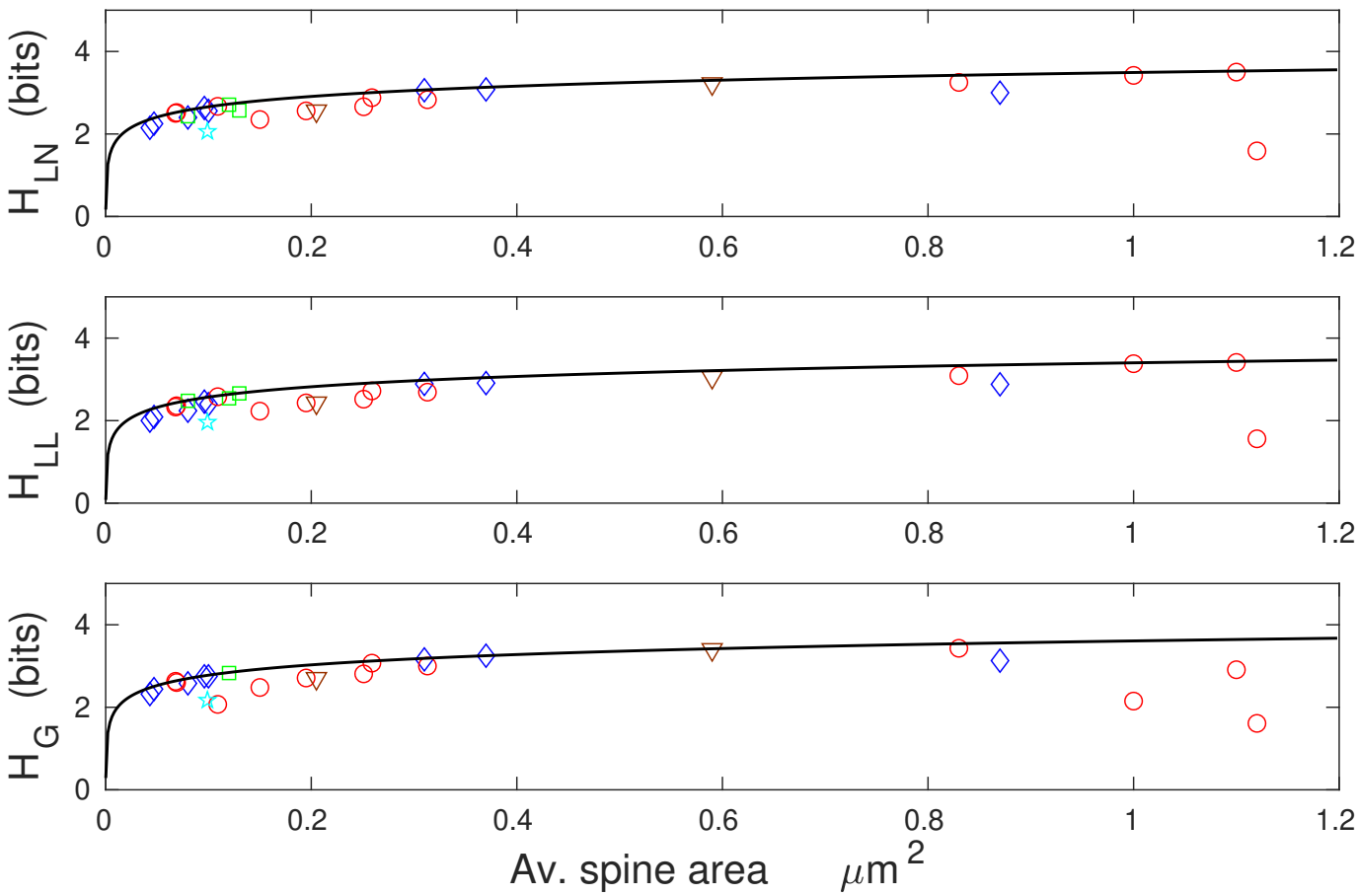

Figure 10

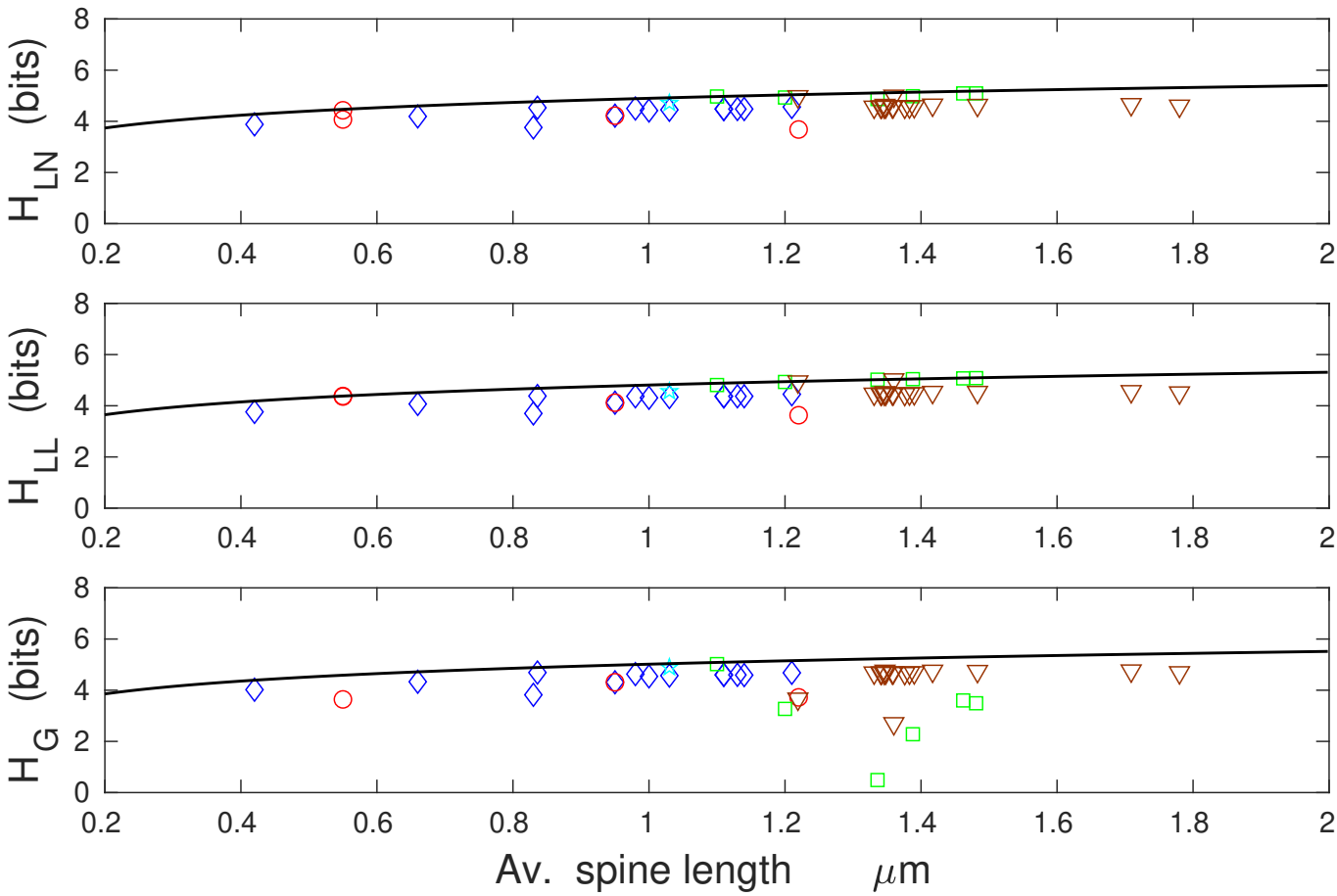
H_LN (bits)
H_LL (bits)
H_G (bits)
0
2
4
6
8
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Av. spine length μm

Figure 11

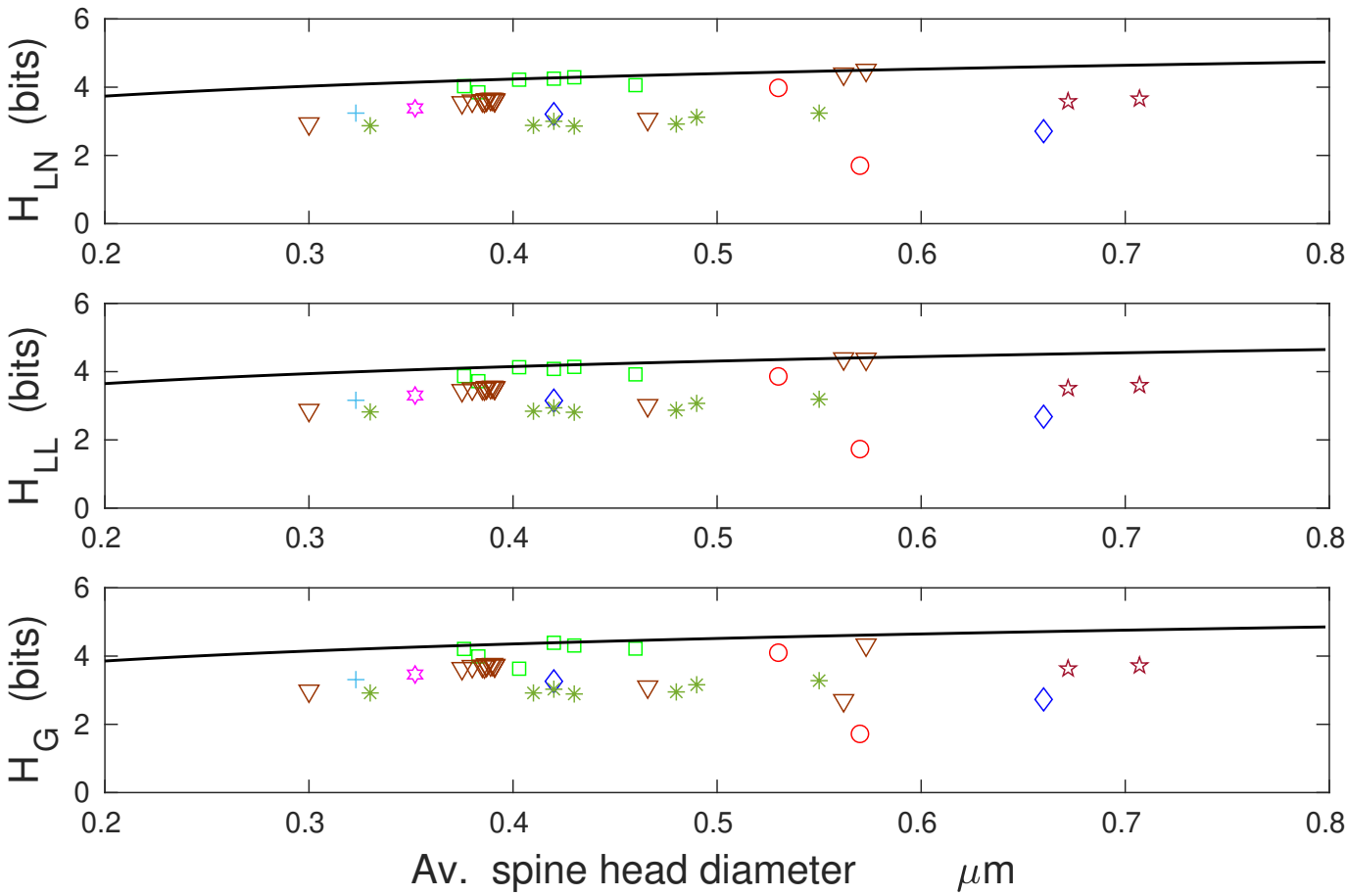
H_LN (bits)
H_LL (bits)
H_G (bits)
0
2
4
6
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Av. spine head diameter μm

Figure 12

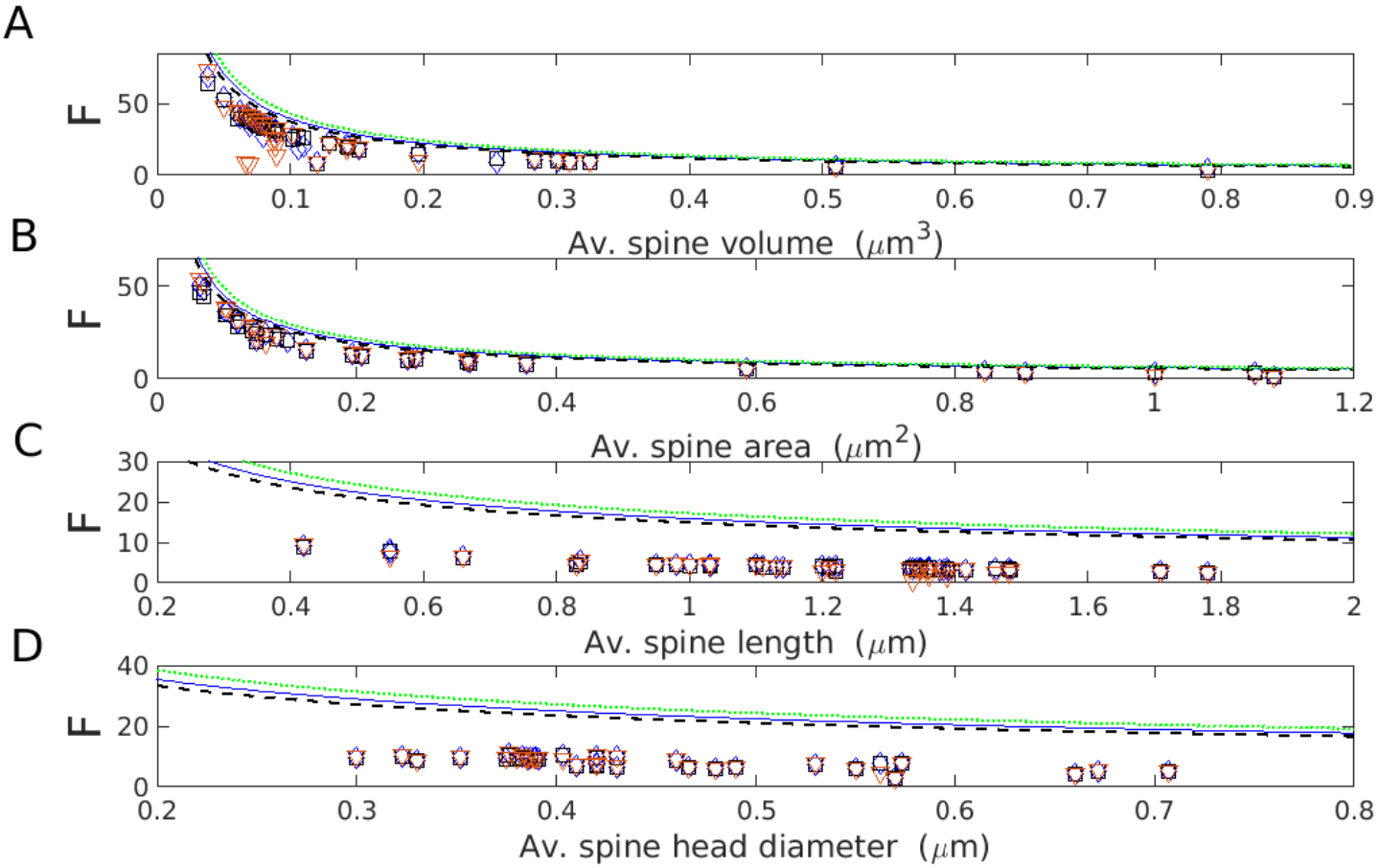
A
F
50
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Av. spine volume (μm^3)
B
F
50
0
0
0.2
0.4
0.6
0.8
1
1.2
Av. spine area (μm^2)
C
F
30
20
10
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Av. spine length (μm)
D
F
40
20
0
0.2
0.3
0.4
0.5
0.6
0.7
0.8
Av. spine head diameter (μm)

Figure 13